(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 130 094 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
05.09.2001 Bulletin 2001/36

(21) Application number: **00114089.6**

(22) Date of filing: **07.07.2000**

(51) Int Cl.[7]: **C12N 15/12**, C12N 15/11, C12N 15/10, C12N 15/70, C12N 15/85, C12N 5/10, C12N 1/21, C07K 14/47, C07K 16/18, C12Q 1/68

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **08.07.1999 JP 19448699**
11.01.2000 JP 2000118774
02.05.2000 JP 2000183765

(71) Applicant: **Helix Research Institute**
Kisarazu-shi, Chiba 292-0812 (JP)

(72) Inventors:
- **Ota, Toshio**
  Fujisawa-shi, Kanagawa 251-0042 (JP)
- **Nishikawa, Tetsuo**
  Tokyo 173-0013 (JP)
- **Isogai, Takao**
  Inashiki-gun, Ibaraki 300-0303 (JP)
- **Hayashi, Koji**
  Ichihara-shi, Chiba 299-0125 (JP)
- **Ishii, Shizuko**
  Kisarazu-shi, Chiba 292-0812 (JP)
- **Kawai, Yuri**
  Kisarazu-shi, Chiba 292-0812 (JP)
- **Wakamatsu, Ai**
  Kisarazu-shi, Chiba 292-0014 (JP)
- **Sugiyama, Tomoyasu**
  Kisarazu-shi, Chiba 292-0045 (JP)
- **Nagai, Keiichi**
  Higashiyamato-shi, Tokyo 207-0022 (JP)
- **Kojima, Shinichi**
  Kisarazu-shi, Chiba 292-0052 (JP)
- **Otsuki, Tetsuji**
  Kisarazu-shi, Chiba 292-0055 (JP)
- **Koga, Hisashi**
  Kisarazu-shi, Chiba 292-0055 (JP)

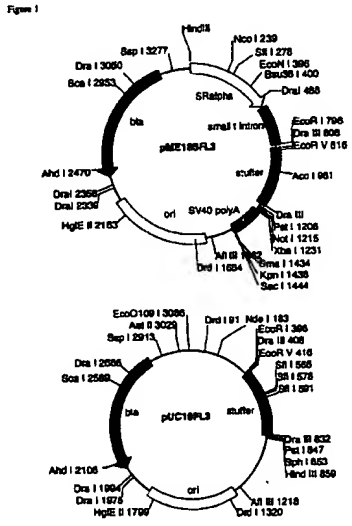
(74) Representative: **VOSSIUS & PARTNER**
Siebertstrasse 4
81675 München (DE)

Remarks:
The sequence listing, which is published as annex to the application documents, was filed after the date of filing. The applicant has declared that it does not include matter which goes beyond the content of the application as filed.

(54) Primers for synthesizing full length cDNA clones and their use

(57) Primers for synthesizing full length cDNAs and their use are provided.

830 cDNA encoding a human protein has been isolated and nucleotide sequences of 5'-, and 3'-ends of the cDNA have been determined. Furthermore, primers for synthesizing the full length cDNA have been provided to clarify the function of the protein encoded by the cDNA. The full length cDNA of the present invention containing the translation start site provides information useful for analyzing the functions of the protein.



EP 1 130 094 A2

Description

FIELD OF THE INVENTION

[0001] The present invention relates to a polynucleotide encoding a novel protein, a protein encoded by the polynucleotide, and new uses of these.

BACKGROUND OF THE INVENTION

[0002] Currently, the sequencing projects, the determination and analysis of the genomic DNA of various living organisms have been in progress all over the world. The whole genomic sequences of more than 10 species of prokaryotes, a lower eukaryote, yeast, and a multicellular eukaryote, C. elegans are already determined. As to human genome, which is supposed to be composed of three thousand million base pairs, the world wide cooperative projects have been under way to analyze it, and the whole structure is predicted to be determined by the years 2002-2003. The aim of the determination of genomic sequence is to reveal the functions of all genes and their regulation and to understand living organisms as a network of interactions between genes, proteins, cells or individuals through deducing the information in a genome, which is a blueprint of the highly complicated living organisms. To understand living organisms by utilizing the genomic information from various species is not only important as an academic subject, but also socially significant from the viewpoint of industrial application.

[0003] However, determination of genomic sequences itself cannnot identify the functions of all genes. For example, as for yeast, only the function of approximately half of the 6000 genes, which is predicted based on the genomic sequence, was able to be deduced. As for human, the number of the genes is predicted to be approximately one hundred thousand. Therefore, it is desirable to establish "a high throughput analysis system of the gene functions" which allows us to identify rapidly and efficiently the functions of vast amounts of the genes obtained by the genomic sequencing.

[0004] Many genes in the eukaryotic genome are split by introns into multiple exons. Thus, it is difficult to predict correctly the structure of encoded protein solely based on genomic information. In contrast, cDNA, which is produced from mRNA that lacks introns, encodes a protein as a single continuous amino acid sequence and allows us to identify the primary structure of the protein easily. In human cDNA research, to date, more than one million ESTs (Expression Sequence Tags) are publicly available, and the ESTs presumably cover not less than 80% of all human genes.

[0005] The information of ESTs is utilized for analyzing the structure of human genome, or for predicting the exon-regions of genomic sequences or their expression profile. However, many human ESTs have been derived from proximal regions to the 3'-end of cDNA, and information around the 5'-end of mRNA is extremely little. Among these human cDNAs, the number of the corresponding mRNAs whose encoding protein sequences are deduced is approximately 7000, and further, the number of full-length therein is only 5500. Thus, even including cDNA registered as EST, the percentage of human cDNA obtained so far is estimated to be 10-15% of all the genes.

[0006] It is possible to identify the transcription start site of mRNA on the genomic sequence based on the 5'-end sequence of a full-length cDNA, and to analyze factors involved in the stability of mRNA that is contained in the cDNA, or in its regulation of expression at the translation stage. Also, since a full-length cDNA contains ATG, the translation start site, in the 5'-region, it can be translated into a protein in a correct frame. Therefore, it is possible to produce a large amount of the protein encoded by the cDNA or to analyze biological activity of the expressed protein by utilizing an appropriate expression system. Thus, analysis of a full-length cDNA provides valuable information which complements the information from genome sequencing. Also, full-length cDNA clones that can be expressed are extremely valuable in empirical analysis of gene function and in industrial application.

[0007] In particular, human secretory proteins or membrane proteins are would be useful by itself as a medicine like tissue plasminogen activator (TPA), or as a target of medicines like membrane receptors. In addition, genes for signal transduction-associated proteins (protein kinases, etc.), glycoprotein-associated proteins, transcription-associated proteins, and disease-associated proteins form a gene group rich in genes whose relationships to human diseases have been elucidated.

[0008] Therefore, it has great significance to isolate novel full-length cDNA clones of human, only few of which has been isolated. Especially, isolation of a novel cDNA clone encoding a secretory protein or membrane protein is desired since the protein itself would be useful as a medicine, and also the clones potentially include a gene associated with diseases. In addition, genes encoding proteins that are associated with signal transduction, glycoprotein, transcription, or diseases are expected to be useful as target molecules for therapy, or as medicines themselves. These genes form a gene group predicted to be strongly associated with diseases. Thus, identification of the full-length cDNA clones encoding those proteins has great significance.

SUMMARY OF THE INVENTION

[0009] An objective of the present invention is to provide a primer that enables synthesizing polynucleotide from human, the resulting polynucleotide or its clone, and a protein encoded by the polynucleotide.

[0010] The inventors have developed a method for efficiently cloning a human full-length cDNA that is predicted by the ATGpr etc. to be a full-length cDNA clone, from a full-length-enriched cDNA library that is synthesized by the oligo-capping method. Then, the inventors determined the nucleotide sequence of the obtained cDNA clones from both 5'- and 3'- ends. By utilizing the sequences, the inventors selected clones that were expected to contain a signal by the PSORT (Nakai K. and Kanehisa M. (1992) Genomics 14: 897-911), and obtained clones that contain a cDNA encoding a secretory protein or membrane protein. Moreover, the inventors specifically selected full-length cDNA clones that encode secretory or membrane proteins, signal transduction-associated proteins, glycoprotein-associated proteins, transcription-associated proteins, or disease-associated proteins from clones homologous to the clones in the Swiss-Prot (http://www.ebi.ac.uk/ebi_docsSwissProt_db/swisshome.html) according to the keywords of SwissProt.

[0011] The full-length cDNA clones of the present invention have high fullness ratio since these were obtained by the combination of (1) construction of a full-length-enriched cDNA library that is synthesized by the oligo-capping method, and (2) a system in which fullness ratio is evaluated from the nucleotide sequence of the 5'-end (in this system, clones are selected based on the estimation by the ATGpr, following the removal of sequences judged not to be full-length when compared with ESTs). However, the primers of the present invention enable obtaining full-length cDNA easily without any special methods mentioned above.

[0012] Homology analysis in which the analysis is carried out against a non-full-length cDNA fragment to postulate the function of a protein encoded by said fragment, is being commonly performed. However, since such analysis is based on the information of the fragment, it is not clear as to whether this fragment corresponds to a part that is functionally important in the protein. In other words, the reliability of the homology analysis based on the information of a fragment is doubtful, as information relating to the structure of the whole protein is not available. However, the homology analysis of the present invention is conducted based on the information of a full-length cDNA comprising the whole coding region of the cDNA, and therefore, the homology of various portions of the protein can be analyzed. Hence, the reliability of the homology analysis has been dramatically improved in the present invention.

[0013] The inventors completed the invention by finding that it is possible to synthesize a novel full-length cDNA by using the combination of a primer that is designed based on the nucleotide sequence of the 5'-ends of the selected full-length cDNA clones and any of an oligo-dT primer or a 3'-primer that is designed based on the nucleotide sequence of the 3'-ends of the selected clones.

[0014] Thus, the present invention relates to primers described below, a method for synthesizing a polynucleotide using the primers, and polynucleotides obtained by the method.

[0015] First, the present invention relates to

(1) use of an oligonucleotide as a primer for synthesizing the polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 1-829 and 2545, or the complementary strand thereof, wherein said oligonucleotide is complementary to said polynucleotide or the complementary strand thereof and comprises at least 15 nucleotides;

(2) a primer set for synthesizing polynucleotides, the primer set comprising an oligo-dT primer and an oligonucleotide complementary to the complementary strand of the polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 1-829 and 2545, wherein said oligonucleotide comprises at least 15 nucleotides; and

(3) A primer set for synthesizing polynucleotides, the primer set comprising a combination of an oligonucleotide comprising a nuclectide sequence complementary to the complementary strand of the polynucleotide comprising a 5'-end nucleotide sequence and an oligonucleotide comprising a nucleotide sequence complementary to the polynucleotide comprising a 3'-end nucleotide sequence, wherein said oligonucleotides comprise at least 15 nucleotides and wherein said combination of 5'-end nucleotide sequence / 3'-end nucleotide sequence is selected from the combinations of 5'-end nucleotide sequence / 3'-end nucleotide sequence set forth in the SEQ ID NOs in Table 1.

[0016] Table 1 shows names of clones obtained in the examples described later, comprising the polynucleotide of the present invention (830 clones), names of nucleotide sequences at the 5'-end and 3'-end of the full-length cDNA, and their corresponding SEQ ID NOs. A blank indicates that the of the 3'-end sequence corresponding to the 5'-end sequence has not been determined the same clone.

[0017] The SEQ ID NO of a 5'-end sequence is shown on the right side of the name of the 5'-end sequence, and the SEQ ID NO of a 3'-end sequence is shown on the right side of the name of the 3'-end sequence.

Table 1

Correspondence between names of clone and the sequence name, and the SEQ ID.				
Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
BNGH41000020	F-BNGH41000020	1		
BNGH41000087	F-BNGH41000087	2		
BNGH41000091	F-BNGH41000091	3		
HEMBA1000006	F-HEMBA1000006	4	R-HEMBA1000006	830
HEMBA1000121	F-HEMBA1000121	5	R-HEMBA1000121	831
HEMBA1000128	F-HEMBA1000128	6	R-HEMBA1000128	832
HEMBA1000275	F-HEMBA1000275	7	R-HEMBA1000275	833
HEMBA1000300	F-HEMBA1000300	8	R-HEMBA1000300	834
HEMBA1000349	F-HEMBA1000349	9	R-nnnnnnnnnnnn	835
HEMBA1000443	F-HEMBA1000443	10		
HEMBA1000462	F-HEMBA1000462	11	R-HEMBA1000462	836
HEMBA1000477	F-HEMBA1000477	12	R-HEMBA1000477	837
HEMBA1000590	F-HEMBA1000590	13	R-HEMBA1000590	838
HEMBA1000634	F-HEMBA1000634	14	R-HEMBA1000634	839
HEMBA1000671	F-HEMBA1000671	15	R-HEMBA1000671	840
HEMBA1000713	F-HEMBA1000713	16	R-HEMBA1000713	841
HEMBA1000732	F-HEMBA1000732	17	R-HEMBA1000732	842
HEMBA1000745	F-HEMBA1000745	18	R-nnnnnnnnnnnn	843
HEMBA1000835	F-HEMBA1000835	19		
HEMBA1000875	F-HEMBA1000875	20	R-HEMBA1000875	844
HEMBA1000907	F-HEMBA1000907	21		
HEMBA1000940	F-HEMBA1000940	22	R-HEMBA1000940	845
HEMBA1000962	F-HEMBA1000962	23	R-HEMBA1000962	846
HEMBA1001184	F-HEMBA1001184	24	R-HEMBA1001184	847
HEMBA1001221	F-HEMBA1001221	25	R-HEMBA1001221	848
HEMBA1001228	F-HEMBA1001228	26	R-HEMBA1001228	849
HEMBA1001272	F-HEMBA1001272	27	R-HEMBA1001272	850
HEMBA1001296	F-HEMBA1001296	28	R-HEMBA1001296	851
HEMBA1001297	F-HEMBA1001297	29	R-HEMBA1001297	852
HEMBA1001390	F-HEMBA1001390	30	R-HEMBA1001390	853
HEMBA1001563	F-HEMBA1001563	31	R-HEMBA1001563	854
HEMBA1001621	F-HEMBA1001621	32	R-HEMBA1001621	855
HEMBA1001878	F-HEMBA1001878	33	R-HEMBA1001878	856
HEMBA1001886	F-HEMBA1001886	34	R-HEMBA1001886	857
HEMBA1002048	F-HEMBA1002048	35	R-HEMBA1002048	858
HEMBA1002131	F-HEMBA1002131	36	R-HEMBA1002131	859
HEMBA1002163	F-HEMBA1002163	37	R-HEMBA1002163	860
HEMBA1002164	F-HEMBA1002164	38		
HEMBA1002167	F-HEMBA1002167	39	R-HEMBA1002167	861
HEMBA1002178	F-HEMBA1002178	40	R-HEMBA1002178	862
HEMBA1002195	F-HEMBA1002195	41	R-HEMBA1002195	863
HEMBA1002227	F-HEMBA1002227	42	R-HEMBA1002227	864
HEMBA1002239	F-HEMBA1002239	43		
HEMBA1002316	F-HEMBA1002316	44	R-HEMBA1002316	865
HEMBA1002420	F-HEMBA1002420	45	R-HEMBA1002420	866
HEMBA1002421	F-HEMBA1002421	46	R-HEMBA1002421	867
HEMBA1002524	F-HEMBA1002524	47	R-HEMBA1002524	868

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.				
Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
NT2RP2002527	F-NT2RP2002527	336	R-NT2RP2002527	1096
NT2RP2002533	F-NT2RP2002533	337	R-NT2RP2002533	1097
NT2RP2002564	F-NT2RP2002564	338	R-NT2RP2002564	1098
NT2RP2002674	F-NT2RP2002674	339	R-NT2RP2002674	1099
NT2RP2002721	F-NT2RP2002721	340	R-NT2RP2002721	1100
NT2RP2002824	F-NT2RP2002824	341	R-NT2RP2002824	1101
NT2RP2002942	F-NT2RP2002942	342	R-NT2RP2002942	1102
NT2RP2002974	F-NT2RP2002974	343	R-NT2RP2002974	1103
NT2RP2002976	F-NT2RP2002976	344	R-NT2RP2002976	1104
NT2RP2003042	F-NT2RP2003042	345	R-NT2RP2003042	1105
NT2RP2003138	F-NT2RP2003138	346		
NT2RP2003179	F-NT2RP2003179	347	R-NT2RP2003179	1106
NT2RP2003210	F-NT2RP2003210	348	R-NT2RP2003210	1107
NT2RP2003302	F-NT2RP2003302	349	R-NT2RP2003302	1108
NT2RP2003369	F-NT2RP2003369	350	R-NT2RP2003369	1109
NT2RP2003383	F-NT2RP2003383	351	R-NT2RP2003383	1110
NT2RP2003390	F-NT2RP2003390	352	R-NT2RP2003390	1111
NT2RP2003469	F-NT2RP2003469	353	R-NT2RP2003469	1112
NT2RP2003545	F-NT2RP2003545	354	R-NT2RP2003545	1113
NT2RP2003593	F-NT2RP2003593	355	R-NT2RP2003593	1114
NT2RP2003599	F-NT2RP2003599	356	R-NT2RP2003599	1115
NT2RP2003655	F-NT2RP2003655	357	R-NT2RP2003655	1116
NT2RP2003664	F-NT2RP2003664	358	R-NT2RP2003664	1117
NT2RP2003931	F-NT2RP2003931	359	R-NT2RP2003931	1118
NT2RP2003940	F-NT2RP2003940	360	R-NT2RP2003940	1119
NT2RP2003950	F-NT2RP2003950	361	R-NT2RP2003950	1120
NT2RP2004069	F-NT2RP2004069	362	R-NT2RP2004069	1121
NT2RP2004108	F-NT2RP2004108	363	R-NT2RP2004108	1122
NT2RP2004141	F-NT2RP2004141	364	R-NT2RP2004141	1123
NT2RP2004179	F-NT2RP2004179	365	R-NT2RP2004179	1124
NT2RP2004205	F-NT2RP2004205	366	R-NT2RP2004205	1125
NT2RP2004447	F-NT2RP2004447	367	R-NT2RP2004447	1126
NT2RP2004495	F-NT2RP2004495	368	R-NT2RP2004495	1127
NT2RP2004524	F-NT2RP2004524	369	R-NT2RP2004524	1128
NT2RP2004556	F-NT2RP2004556	370	R-NT2RP2004556	1129
NT2RP2004606	F-NT2RP2004606	371	R-NT2RP2004606	1130
NT2RP2004648	F-NT2RP2004648	372	R-NT2RP2004648	1131
NT2RP2004670	F-NT2RP200467()	373	R-NT2RP2004670	1132
NT2RP2004794	F-NT2RP2004794	374	R-NT2RP2004794	1133
NT2RP2004837	F-NT2RP2004837	375	R-NT2RP2004837	1134
NT2RP2004847	F-NT2RP2004847	376	R-NT2RP2004847	1135
NT2RP2005027	F-NT2RP2005027	377	R-NT2RP2005027	1136
NT2RP2005069	F-NT2RP2005069	378	R-NT2RP2005069	1137
NT2RP2005163	F-NT2RP2005163	379	R-NT2RP2005163	1138
NT2RP2005181	F-NT2RP2005181	380	R-NT2RP2005181	1139
NT2RP2005247	F-NT2RP2005247	381	R-NT2RP2005247	1140
NT2RP2005378	F-NT2RP2005378	382	R-NT2RP2005378	1141
NT2RP2005391	F-NT2RP2005391	383	R-NT2RP2005391	1142

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.				
Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
Y79AA1001795	F-Y79AA1001795	816	R-Y79AA1001795	1560
Y79AA1001799	F-Y79AA1001799	817	R-Y79AA1001799	1561
Y79AA1001803	F-Y79AA1001803	818	R-Y79AA1001803	1562
Y79AA1001863	F-Y79AA1001863	819	R-Y79AA1001863	1563
Y79AA1002022	F-Y79AA1002022	820	R-Y79AA1002022	1564
Y79AA1002058	F-Y79AA1002058	821		
Y79AA1002121	F-Y79AA1002121	822	R-nnnnnnnnnnn	1565
Y79AA1002129	F-Y79AA1002129	823	R-nnnnnnnnnnn	1566
Y79AA1002213	F-Y79AA1002213	824	R-Y79AA1002213	1567
Y79AA1002334	F-Y79AA1002334	825	R-Y79AA1002334	1568
Y79AA1002373	F-Y79AA1002373	826	R-Y79AA1002373	1569
Y79AA1002376	F-Y79AA1002376	827	R-Y79AA1002376	1570
Y79AA1002378	F-Y79AA1002378	828	R-Y79AA1002378	1571
Y79AA1002381	F-Y79AA1002381	829	R-Y79AA1002381	1572
NT2RP2006580	F-NT2RP2006580	2545	R-NT2RP2006580	2546

The sequence name starting from "F" means the name of 5'-end sequence, and the sequence name starting from "R" means the name of 3'-end sequence. A blank indicates that the 3'-end sequence corresponding to the 5'-end sequence has not been determined in the clone.

[0018] Furthermore, the present invention relates to the use of the above primers, as described below.

(4) A polynucleotide which can be synthesized with the primer set of (2) or (3).
(5) A polynucleotide comprising a coding region in the polynucleotide of (4).
(6) A substantially pure protein encoded by polynucleotide of (4).
(7) A partial peptide of the protein of (6).

[0019] In addition, the present invention comprises a polynucleotide described below and a protein encoded by the polynucleotide.

(8) An isolated polynucleotide selected from the group consisting of

(a) a polynucleotide comprising a coding region of the nucleotide sequence set forth in any one of the SEQ ID NOs in Table 370;
(b) a polynucleotide comprising a nucleotide sequence encoding a protein comprising the amino acid sequence set forth in any one of the SEQ ID NOs in Table 370;
(c) a polynucleotide comprising a nucleotide sequence encoding a protein comprising an amino acid sequence selected from the amino acid sequences set forth in the SEQ ID NOs in Table 370, in which one or more amino acids are substituted, deleted, inserted, and/or added, wherein said protein is functionally equivalent to the protein comprising said amino acid sequence selected from the amino acid sequences set forth in the SEQ ID NOs in Table 370;
(d) a polynucleotide that hybridizes with a polynucleotide comprising a nucleotide sequence selected from the nucleotide sequences set forth in the SEQ ID NOs in Table 370, and that comprises a nucleotide sequence encoding a protein functionally equivalent to the protein encoded by the nucleotide sequence selected from the nucleotide sequences set forth in the SEQ ID NOs in Table 370;
(e) a polynucleotide comprising a nucleotide sequence encoding a partial amino acid sequence of a protein encoded by the polynucleotide of (a) to (d);
(f) a polynucleotide comprising a nucleotide sequence with at least 70% identity to the nucleotide sequence set forth in any one of the SEQ ID NOs in Table 370.

(9). A substantially pure protein encoded by the polynucleotide of (8).
(10) An antibody against the protein or peptide of any one of (6), (7), and (9).
(11) A vector comprising the polynucleotide of (5) or (8).

(12) A transformant carrying the polynucleotide of (5) or (8), or the vector of (11).
(13) A transformant expressively carrying the polynucleotide of (5) or (8), or the vector of (11).
(14) A method for producing the protein or peptide of any one of (6), (7), and (9), comprising culturing the transformant of (13) and recovering the expression product.
(15) An oligonucleotide comprising; the nucleotide sequence set forth in any one of the SEQ ID NOs in Table 370 or the nucleotide sequence complementary to the complementary strand thereof, wherein said oligonucleotide comprises 15 nucleotides or more.
(16) Use of the oligonucleotide of (15) as a primer for synthesizing a polynucleotide.
(17) Use of the oligonucleotide of (15) as a probe for detecting a gene.
(18) An antisense polynucleotide against the polynucleotide of (8), or the portion thereof.
(19) A method for synthesizing a polynucleotide, the method comprising:

a) synthesizing a complementary strand using a cDNA library as a template, and using the primer set of (2) or (3), or the primer of (16); and
b) recovering the synthesized product.

(20) The method of (19), wherein the cDNA library is obtainable by oligo-capping method.
(21) The method of (19), wherein the complementary strand is obtainable by PCR.
(22) A method for detecting the polynucleotide of (8), the method comprising:

a) incubating a target polynucleotide with the oligonucleotide of (15) under the conditions where hybridization occurs, and
b) detecting the hybridization of the target polynucleotide with the oligonucleotide of (15).

(23) A database of polynucleotides and/or proteins, the database comprising information on at least one sequence selected from the nucleotide sequences set forth in the SEQ ID NOs in Table 370 and/or the amino acid sequences set forth in the SEQ ID NOs in Table 370, or a medium on which the database is stored.

[0020] Any patents, patent applications, and publications cited herein are incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021]

Figure 1 shows the restriction maps of vectors pME18SFL3 and pUC19FL3.
Figure 2 shows the reproducibility of gene expression analysis. The ordinate and the abscissa show the intensities of gene expression obtained in experiments different from each other.
Figure 3 shows the detection limit in gene expression analysis. The intensity of expression is shown in the ordinate, and the concentration (μg/ml) of the probe used is shown in the abscissa.

DETAILED DESCRIPTION OF THE INVENTION

[0022] Herein, "polynucleotide" is defined as a molecule in which multiple nucleotides are polymerized. There are no limitations in the number of the polymerized nucleotides. In case that the polymer contains relatively low number of nucleotides, it is also described as an "oligonucleotide". The polynucleotide or the oligonucleotide of the present invention can be a natural or chemically synthesized product. Alternatively, it can be synthesized using a template DNA by an enzymatic reaction such as PCR.
[0023] All the cDNA provided by the invention are full-length cDNA. Herein, a "full-length cDNA" is defined as a cDNA which contains both ATG codon (the translation start site) and the stop codon. Accordingly, the untranslated regions, which are originally found in the upstream or downstream of the protein coding region in natural mRNA, may or may not be contained.
An "isolated polynucleotide" is a polynucleotide the structure of which is not identical to that of any naturally occurring nucleic acid or to that of any fragment of a naturally occurring genomic nucleic acid spanning more than three separate genes. The term therefore covers, for example,

(a) a DNA which has the sequence of part of a naturally occurring genomic DNA molecule but is not flanked by both of the coding sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs;

(b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA;
(c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and
(d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Specifically excluded from this definition are nucleic acids present in mixtures of different (i) DNA molecules, (ii) transfected cells, or (iii) cell clones: e.g., as these occur in a DNA library such as a cDNA or genomic DNA library.

[0024] The term "substantially pure" as used herein in reference to a given polypeptide means that the protein or polypeptide is substantially free from other biological macromolecules. The substantially pure protein or polypeptide is at least 75% (e.g., at least 80, 85, 95, or 99%) pure by dry weight. Purity can be measured by any appropriate standard method, for example, by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

[0025] All the clones of the present invention (830 clones) are novel and covering full-length, and also predicted to encode any of the following functional protein:

secretory proteins,
membrane proteins,
proteins associated to signal transduction (signal transduction-associated proteins; e.g. protein kinases, etc.),
proteins associated to a glycoprotein (glycoprotein-associated proteins),
proteins associated with transcription (transcription-associated proteins),
proteins associated with diseases (disease-associated proteins),
or, enzymes and/or metabolism-associated proteins, cell division- and/or cell proliferation-associated proteins, cytoskeleton-associated proteins, nuclear proteins, DNA-and/or RNA-binding proteins, ATP- and/or GTP-binding proteins, protein synthesis- and/or protein transport-associated proteins, and cellular defense-associated proteins.

[0026] Furthermore, all the cDNA clones of the present invention can be characterized as follows:

(1) a cDNA that is obtained by the oligo-capping method, which provides cDNA with high fullness ratio. The cDNA was selected by the score in the ATGpr (described as ATGpr1, as well), which is a program for prediction of the fullness of the 5'-end of cDNA based on the features of the 5'-end sequence. In addition, the PSORT, which is a program for prediction of the existence of the signal sequence selected, cDNA that contains a signal sequence in the 5'-end, or transmembrane region in the protein coding region. Furthermore, the homology search with the 5'-end sequences confirmed that, the selected clones were not identical to any of the known human mRNA (namely novel);
or,
(2) a cDNA that is obtained by the oligo-capping method, which provides cDNA with high fullness ratio. The cDNA was selected by the score in the ATGpr, which is a program for prediction of the fullness of the 5'-end based on the features of the 5'-end sequence. Furthermore, the a cDNA that has relative homology with an amino acid sequence of a protein with known functions was selected by the BLAST search (Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J. (1990) J. Mol. Biol. 215: 403-410 ; Gish W., and States D.J. (1993) Nature Genet. 3: 266-272) on the SwissProt database using the 5'-end sequence. In addition, the homology search using the 5'-end sequence confirmed that the selected clones were not identical to any of the known human mRNA (namely novel).

[0027] All clones are obtainable as a full-length clone by such a method as PCR (Current Protocols in Molecular Biology, Ausubel et al. edit, (1987) John Wiley & Sons, Section 6.1-6.4) using both the 5'- and 3'-end sequences, or using the 5'-end sequence and an oligo-dT primer that corresponds to the polyA sequence.

[0028] Specifically, PCR can be performed using an oligonucleotide that has 15 nucleotides longer, and specifically hybridizes with the complementary strand of the polynucleotide that contains the nucleotide sequence selected from the 5'-end sequences shown in Table 1 (SEQ ID NO: 1-829, and SEQ ID NO: 2545), and an oligo-dT primer as a 5'-, and 3'-primer, respectively. The length of the primers is usually 15-100 bp, and favorably between 15-35 bp. In case of LA PCR, which is described below, the primer length of 25-35 bp may provide a good result.

[0029] A method to design a primer that enables a specific amplification based on the given nucleotide sequence is known to those skilled in the art (Current Protocols in Molecular Biology, Ausubel et al. edit, (1987) John Wiley & Sons, Section 6.1-6.4). In designing a primer based on the 5'-end sequence, the primer is designed so as that, in principle, the amplification products will include the translation start site. Accordingly, in case that a given 5'-end nucleotide sequence is the 5'-untranslated region (5'UTR), any part of the sequence can be used as a 5'-primer as far as the specificity toward the target cDNA is insured. The translation start site can be predicted using a known method such

as the ATGpr as described below.

[0030] When synthesizing a polynucleotide, the target nucleotide sequence to be amplified can extend to several thousand bp in some cDNA. However, it is possible to amplify such a long nucleotides by using such as LA PCR (Long and Accurate PCR). It is advantageous to use LA PCR when synthesizing long DNA In LA PCR, in which a special DNA polymerase having 3' □5' exonuclease activity is used, misincorporated nucleotides can be removed. Accordingly, accurate synthesis of the complementary strand can be achieved even with a long nucleotide sequence. By using LA PCR, it is reported that amplification of a nucleotide with 20 kb longer can be achieved under desirable condition (Takeshi Hayashi (1996) Jikken-Igaku Bessatsu, "Advanced Technologies in PCR" Youdo-sha).

[0031] A template DNA for synthesizing the cDNA of the present invention can be obtained by using cDNA libraries that are prepared by various methods. The full-length cDNA clones obtained here are those with high fullness ratio, which were obtained using a combination of (1) a method to prepare a full-length-enriched cDNA library using the oligo-capping method, and (2) an estimation system for fullness using the 5'-end sequence (selection based on the estimation by the ATGpr after removing clones that are non-full-length compared to the ESTs). However, it is possible to easily obtain a full-length cDNA by using the primers that are provided by the present invention, not by the above described specialized method.

[0032] The problem with the cDNA libraries prepared by the known methods or commercially available is that mRNA contained in the libraries has very low fullness ratio. Thus, it is difficult to screen full-length cDNA clone directly from the library using ordinary cloning methods. The present invention has revealed a primer that is capable of synthesizing a full-length cDNA. If provided with primers, it is possible to synthesize a target full-length cDNA by using enzymatic reactions such as PCR. In particular, a full-length-enriched cDNA library, synthesized by methods such as oligo-capping, is desirable to synthesize a full-length cDNA with more reliability.

[0033] Once the nucleotide sequences of the full-length cDNAs obtained in the present invention is determined, it is possible to predict the functions of the proteins encoded by the cDNA clones, for example, by searching the databases such as GenBank (http://www.ncbi.nlm.nih.gov/web/GenBank/), Swiss-Prot (http://www.ebi.ac.uk/ebi_docsSwiss-Prot_db/swisshome.html), UniGene (http://www.ncbi.nlm.nih.gov/UniGene) for homologies of the cDNAs, or by searching the amino acid sequences deduced from the full-length nucleotide sequences for signal sequence by using software such as PSORT (K. Nakai & M. Kanehisa, Genomics, 14: 897-991 (1992), for transmembrane region by using software such as SOSUI (T. Hirokawa et al., Bioinformatics, 14:378-379 (1998); Mitsui Knowledge Industry Co., Ltd.) or for motif by using software such as Pfam (http://www.sanger.ac.uk/Software/Pfam/index.shtml) or PROSITE (http://www.expasy.ch/prosite). As a matter of course, the functions are often predictable by using partial sequence information (preferably 300 nucleotides or more) instead of the full-length nucleotide sequences. However, the result of the prediction obtained by using partial sequence information does not always agree with the result obtained by using full-length nucleotide sequence, and thus it is needless to say that the prediction of function is preferably performed based on the full-length nucleotide sequences.

[0034] Homology search using each of GenBank, Swiss-Prot and UniGene was performed for the 826 clones whose full-length nucleotide sequences had been determined (HEMBA1005337, NT2RM1000407, NT2RM2001767, and NT2RP3003939 are not full-length). The amino acid sequences deduced from the full-length nucleotide sequences were searched for functional domains by using analytical software programs, PSORT, SOSUI and Pfam. Based on the results, proteins encoded by the cDNA clones were grouped into some categories and their functions were predicted.

[0035] The following 437 clones were categorized into secretory and/or membrane proteins. The clones categorized into secretory and/or membrane proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "growth factor", "cytokine", "hormone", "signal", "transmembrane", "membrane", "extracellular matrix", "receptor", "G-protein coupled receptor", "ionic channel", "voltage-gated channel", "calcium channel", "cell adhesion", "collagen" or "connective tissue"; those which matched the data, suggesting that the proteins are secretory and/or membrane proteins; or those which matched with the full-length sequences of GenBank or UniGene database similar description; and, further, those predicted to have an N-terminal signal sequence or a transmembrane region as a result of domain search for the amino acid sequences deduced from the full-length nucleotide sequences.

BNGH41000020, BNGH41000087, BNGH41000091, HEMBA1000121, HEMBA1000128, HEMBA1000349, HEMBA1000477, HEMBA1000590, HEMBA1000713, HEMBA1000732, HEMBA1000745, HEMBA1000835, HEMBA1000940, HEMBA1000962, HEMBA1001221, HEMBA1001228, HEMBA1001621, HEMBA1002131, HEMBA1002163, HEMBA1002167, HEMBA1002178, HEMBA1002195, HEMBA1002227, HEMBA1002420, HEMBA1002421, HEMBA1002767, HEMBA1003047, HEMBA1003101, HEMBA1003230, HEMBA1003392, HEMBA1003530, HEMBA1003602, HEMBA1003732, HEMBA1003945, HEMBA1004110, HEMBA1004250, HEMBA1004391, HEMBA1004444, HEMBA1004454, HEMBA1004505, HEMBA1004797, HEMBA1004982, HEMBA1005070, HEMBA1005449, HEMBA1005522, HEMBA1005545, HEMBA1005698, HEMBA1005945, HEMBA1006171, HEMBA1006299, HEMBA1006311, HEMBA1006335, HEMBA1006357, HEMBA1006430, HEMBA1006482, HEMBA1006707, HEMBA1006724, HEMBA1006749, HEMBA1006902, HEMBA1006960, HEMBA1007241, HEMBB1000407, HEMBB1000447, HEMBB1000567, HEMBB1000679, HEMBB1000881,

THYRO1001478, THYRO1001523, THYRO1001529, THYRO1001641, THYRO1001702, THYRO1001725, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000426, Y79AA1000521, Y79AA1000876, Y79AA1000888, Y79AA1000959, Y79AA1001013, Y79AA1001212, Y79AA1001264, Y79AA1001328, Y79AA1001426, Y79AA1001427, Y79AA1001430, Y79AA1001727, Y79AA1001787, Y79AA1001795, Y79AA1001799, Y79AA1001803, Y79AA1002022, Y79AA1002058, Y79AA1002129, Y79AA1002213, Y79AA1002373,

[0036] The following 146 clones were categorized into glycoprotein-associated proteins. The clones categorized into glycoprotein-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keyword "glycoprotein"; those which matched the data suggesting that the proteins are glycoprotein; or those which matched the full-length sequences of GenBank or UniGene database with similar description.

BNGH41000087, BNGH41000091, HEMBA1000349, HEMBA1000590, HEMBA1000745, HEMBA1000835, HEMBA1001221, HEMBA1001228, HEMBA1001621, HEMBA1002131, HEMBA1002178, HEMBA1002421, HEMBA1002767, HEMBA1003230, HEMBA1003392, HEMBA1004250, HEMBA1004391, HEMBA1004444, HEMBA1004505, HEMBA1005449, HEMBA1005522, HEMBA1005545, HEMBA1006707, HEMBA1006749, HEMBA1006902, HEMBB1000679, HEMBB1000881, HEMBB1001048, HEMBB1002120, HEMBB1002245, HEMBB1002427, MAMMA1000102, MAMMA1000591, MAMMA1000681, MAMMA1001043, MAMMA1001237, MAMMA1002070, MAMMA1002586, MAMMA1003126, NT2RM1000462, NT2RM1000580, NT2RM2001792, NT2RM2001818, NT2RM2001939, NT2RM2001941, NT2RM4000198, NT2RM4000284, NT2RM4000417, NT2RM4000648, NT2RM4000997, NT2RM4001325, NT2RM4002352, NT2RP1000613, NT2RP1000981, NT2RP1001004, NT2RP2000616, NT2RP2000694, NT2RP2000903, NT2RP2001480, NT2RP2001755, NT2RP2002533, NT2RP2003042, NT2RP2003210, NT2RP2004205, NT2RP2004606, NT2RP2005027, NT2RP2005181, NT2RP2005541, NT2RP2005597, NT2RP2005883, NT2RP2006004, NT2RP2006042, NT2RP2006269, NT2RP3000304, NT2RP3000616, NT2RP3000921, NT2RP3001650, NT2RP3002160, NT2RP3002737, NT2RP3002958, NT2RP3003000, NT2RP3003532, NT2RP3004130, NT2RP3004133, NT2RP3004481, NT2RP3004552, NT2RP3004640, NT2RP4000108, NT2RP4001467, NT2RP4002750, OVARC1000003, OVARC1000553, OVARC1000811, OVARC1000873, OVARC1001336, OVARC1001607, OVARC1001991, PLACE1000033, PLACE1000740, PLACE1001016, PLACE1001123, PLACE1001231, PLACE1001464, PLACE1001655, PLACE1001836, PLACE1002355, PLACE1002374, PLACE1002905, PLACE1002911, PLACE1003573, PLACE1003737, PLACE1003772, PLACE1003839, PLACE1004282, PLACE1004441, PLACE1004450, PLACE1004520, PLACE1004648, PLACE1005003, PLACE1005426, PLACE1006071, PLACE1006073, PLACE1006290, PLACE1007081, PLACE1007845, PLACE1008716, PLACE1008744, PLACE1008985, PLACE1010251, PLACE1010784, PLACE1010968, PLACE1011116, PLACE3000181, PLACE3000213, PLACE4000354, THYRO1000036, THYRO1000196, THYRO1000584, THYRO1000956, THYRO1001266, Y79AA1000270, Y79AA1000426, Y79AA1001727, Y79AA1001795, Y79AA1002022, Y79AA1002213,

[0037] The following 57 clones were categorized into signal transduction-associated proteins. The clones categorized into signal transduction-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "serine/threonine-protein kinase", "tyrosine-protein kinase" or "SH3 domain"; those which matched the data suggesting that the proteins are signal transduction-associated proteins (for example, "ADP-ribosylation factor"); or those which matched the full-length sequences of GenBank or UniGene database with similar description; and, further, those which was similarly predicted to be signal transduction-associated proteins based on the matching data of Pfam.

HEMBA1000006, HEMBA1002195, HEMBA1002227, HEMBA1002551, HEMBA1005084, HEMBA1005929, HEMBA1006658, HEMBA1006916, MAMMA1000881, MAMMA1001150, MAMMA1001310, MAMMA1002142, NT2RM2001902, NT2RP1001020, NT2RP1001031, NT2RP2001469, NT2RP2001529, NT2RP2001769, NT2RP2003179, NT2RP2003545, NT2RP2004670, NT2RP3000011, NT2RP3000022, NT2RP3000172, NT2RP3000201, NT2RP3000820, NT2RP3003527, NT2RP3003849, NT2RP3003874, NT2RP3004067, NT2RP4000634, NT2RP4000962, OVARC1000255, OVARC1000529, OVARC1000916, OVARC1001338, OVARC1001569, PLACE1002329, PLACE1003135, PLACE1003598, PLACE1005519, PLACE1006208, PLACE1008282, PLACE1008297, PLACE1010081, PLACE1011364, PLACE1011824, THYRO1001457, THYRO1001593, THYRO1001700, THYRO1001770, Y79AA1000777, Y79AA1000967, Y79AA1002376, Y79AA1002381, HEMBB1000668, NT2RM4001377

[0038] The following 81 clones were categorized into transcription-associated proteins. The clones categorized into transcription-associated proteins are those which keywords "transcription regulation", "zinc finger" or "homeobox" matched the full-length sequences of Swiss-Prot database; those which the matched the data suggesting that the proteins were transcription-associated proteins; or those which matched the full-length sequences of GenBank or UniGene database with similar description; and, further, those which was similarly predicted to be transcription-associated proteins based on the matching data of Pfam.

PLACE1001183 PLACE1001229 PLACE1001407
PLACE1001536 PLACE1001788 PLACE1002080
PLACE1002095 PLACE1002374 PLACE1002518
PLACE1003407 PLACE1003428 PLACE1003460
PLACE1003839 PLACE1003845 PLACE1004028
PLACE1004199 PLACE1004282 PLACE1004305
PLACE1004482 PLACE1004637 PLACE1005005
PLACE1005250 PLACE1005383 PLACE1005410
PLACE1005544 PLACE1005569 PLACE1005601
PLACE1005660 PLACE1005669 PLACE1005725
PLACE1005768 PLACE1005927 PLACE1006079
PLACE1006093 PLACE1006219 PLACE1006277
PLACE1006443 PLACE1006786 PLACE1006809
PLACE1007040 PLACE1007096 PLACE1007296
PLACE1007626 PLACE1007971 PLACE1008469
PLACE1008984 PLACE1008985 PLACE1009067
PLACE1009196 PLACE1009527 PLACE1009982
PLACE1010251 PLACE1011236 PLACE2000219
PLACE4000455 SKNMC1000004 SKNMC1000014
THYRO1000036 THYRO1000099 THYRO1000196
THYRO1000795 THYRO1000999 THYRO1001237
THYRO1001327 THYRO1001478 THYRO1001495
THYRO1001523 THYRO1001702 THYRO1001725
Y79AA1000226 Y79AA1000270 Y79AA1000426
Y79AA1000521 Y79AA1000776 Y79AA1000959
Y79AA1001013 Y79AA1001056 Y79AA1001264
Y79AA1001328 Y79AA1001427 Y79AA1001430
Y79AA1001530 Y79AA1001592 Y79AA1001793
Y79AA1001795 Y79M1001803 Y79AA1001863
Y79AA1002022 Y79AA1002373

[0056] In the example mentioned below, the 254 clones as described above were categorized into three groups according to their maximal value in the ATGpr and the result in the PSORT, which are shown in Table 7-10, 11, 12 (246 clones), and Table 13, 14, 15 (8 clones). In the tables, the name of clone, indicate the name of the clone that was selected by the ATGpr and the PSORT; the name of sequence indicates the name of the 5'-end sequence of the clone on the left; the maximal ATGpr score indicates the maximal ATGpr1 score of the 5'-end sequence shown on the left; and signal indicates the presence of the signal sequence according to the prediction by the PSORT. In addition, the representative sequence is the sequence that has the longest sequence among the cluster in which the 5'-end sequence on the left was included. The maximal ATGpr score and signal on the right indicate the maximal ATGpr1 score of the representative sequence, and the presence of a signal sequence in the representative sequence according to the prediction by the PSORT, respectively. The 170 clones shown in Table 7-10, having the maximal score in the ATGpr1 higher than 0.5, and predicted to possess a signal sequence by the PSORT, are very likely to be full-length and encode a secretory or membrane protein. The 35 clones in Table 11, which have the maximal score in the ATGpr1 0.3 or higher and less than 0.5, and predicted to have a signal sequence, are also as well. And, the 41 clones in Table 12, having the maximal score in the ATGpr1 0 or higher and less than 0.3, and predicted to have a signal sequence, are likely to be full-length and encode a secretory or membrane protein.

[0057] The 8 clones in Table 13 (4 clones), Table 14 (2 clones), and Table 15 (2 clones) have the maximal score in the ATGpr1 0.5 or higher, 0.3 or higher and less than 0.5, and 0 or higher and less than 0.3, respectively, and are predicted to have no signal sequence by the PSORT. However, these clones contain a region that is recognized by the PSORT to be a signal sequence within the representative sequence composing the same cluster. Thus, the clones were judged as a full-length clone which encodes a membrane protein, especially.

[0058] The clones selected by the score in the ATGpr and by the keywords in the top hit data in the SwissProt are likely to encode a secretory or membrane protein, or proteins with functions associated to signal transduction, glycoprotein, transcription, and diseases according to the respective keywords. These 659 clones are shown below. Here, top hit data is defined to be data of known amino acid sequence which is identified to be the most homologous sequence in homology search using the SwissProt.

BNGH41000020 BNGH41000087 BNGH41000091

NT2RM2000497 NT2RM2000565 NT2RM2000582
NT2RM2000589 NT2RM2000622 NT2RM2000632
NT2RM2000773 NT2RM2001126 NT2RM2001558
NT2RM2001626 NT2RM2001738 NT2RM2001767
NT2RM2001792 NT2RM2001818 NT2RM2001902
NT2RM2001939 NT2RM2001941 NT2RM4000100
NT2RM4000198 NT2RM4000284 NT2RM4000295
NT2RM4000326 NT2RM4000417 NT2RM4000444
NT2RM4000587 NT2RM4000593 NT2RM4000648
NT2RM4000761 NT2RM4000965 NT2RM4001377
NT2RM4001735 NT2RM4001843 NT2RM4002352
NT2RP1000002 NT2RP1000050 NT2RP1000181
NT2RP1000239 NT2RP1000261 NT2RP1000271
NT2RP1000300 NT2RP1000325 NT2RP1000465
NT2RP1000468 NT2RP1000551 NT2RP1000579
NT2RP1000613 NT2RP1000679 NT2RP1000740
NT2RP1000981 NT2RP1001004 NT2RP1001020
NT2RP1001031 NT2RP2000092 NT2RP2000178
NT2RP2000240 NT2RP2000394 NT2RP2000447
NT2RP2000514 NT2RP2000533 NT2RP2000610
NT2RP2000616 NT2RP2000649 NT2RP2000663
NT2RP2000694 NT2RP2000712 NT2RP2000739
NT2RP2000818 NT2RP2000903 NT2RP2001200
NT2RP2001223 NT2RP2001276 NT2RP2001388
NT2RP2001469 NT2RP2001480 NT2RP2001495
NT2RP2001514 NT2RP2001529 NT2RP2001538
NT2RP2001562 NT2RP2001662 NT2RP2001755
NT2RP2001769 NT2RP2001817 NT2RP2001878
NT2RP2001903 NT2RP2001921 NT2RP2001948
NT2RP2001956 NT2RP2002063 NT2RP2002188
NT2RP2002232 NT2RP2002304 NT2RP2002409
NT2RP2002510 NT2RP2002527 NT2RP2002533
NT2RP2002564 NT2RP2002824 NT2RP2002942
NT2RP2002974 NT2RP2002976 NT2RP2003042
NT2RP2003138 NT2RP2003179 NT2RP2003210
NT2RP2003302 NT2RP2003369 NT2RP2003390
NT2RP2003469 NT2RP2003545 NT2RP2003593
NT2RP2003655 NT2RP2003664 NT2RP2003931
NT2RP2003940 NT2RP2003950 NT2RP2004069
NT2RP2004108 NT2RP2004141 NT2RP2004205
NT2RP2004447 NT2RP2004606 NT2RP2004648
NT2RP2004670 NT2RP2004794 NT2RP2004847
NT2RP2005069 NT2RP2005163 NT2RP2005181
NT2RP2005247 NT2RP2005378 NT2RP2005391
NT2RP2005425 NT2RP2005535 NT2RP2005541
NT2RP2005597 NT2RP2005632 NT2RP2005666
NT2RP2005774 NT2RP2005878 NT2RP2005883
NT2RP2005941 NT2RP2005994 NT2RP2006004
NT2RP2006042 NT2RP2006092 NT2RP2006099
NT2RP2006134 NT2RP2006269 NT2RP2006512
NT2RP3000011 NT2RP3000022 NT2RP3000059
NT2RP3000063 NT2RP3000125 NT2RP3000148
NT2RP3000171 NT2RP3000172 NT2RP3000201
NT2RP3000232 NT2RP3000304 NT2RP3000378
NT2RP3000427 NT2RP3000436 NT2RP3000444
NT2RP3000481 NT2RP3000616 NT2RP3000645
NT2RP3000652 NT2RP3000676 NT2RP3000677
NT2RP3000721 NT2RP3000820 NT2RP3000838

HEMBA1007013 NT2RP2003655 PLACE1005544
HEMBB1000447 NT2RP2003664 PLACE1005569
HEMBB1000567 NT2RP2004447 PLACE1005927
HEMBB1001200 NT2RP2006042 PLACE1006277
HEMBB1002427 NT2RP2006269 PLACE1006443
MAMMA1000591 NT2RP3000481 PLACE1007096
MAMMA1000681 NT2RP3000645 PLACE1007296
MAMMA1001043 NT2RP3001012 PLACE1008469
MAMMA1001893 NT2RP3001195 PLACE1008984
MAMMA1001957 NT2RP3001560 PLACE1008985
MAMMA1002070 NT2RP3002160 PLACE1009527
MAMMA1002165 NT2RP3002836 PLACE1010251
MAMMA1002633 NT2RP3002958 PLACE1011236
NT2RP3003076 SKNMC1000014
NT2RP3003354 THYPO1000196
NT2RP3004133 THYRO1000795
NT2RP3004309 THYRO1001478
NT2RP4001879 THYRO1001702
NT2RP4002451 Y79AA1000270
OVARC1000439 Y79AA1000426
OVARC1001222 Y79AA1001803
Y79AA1001863

[0060] The 446 clones in Table 16, 17, 18, 19, and 20, and NT2RP2006580 are predicted to encode a secretory or membrane protein. Among them, 77 clones were identical to the clones selected by the score in the ATGpr and the prediction by the PSORT for the existence of a signal sequence (overlapping with any of the 254 clones listed in Table 7-15). Besides, many clones were turned out to be identical to the clones selected as a protein associated with a glycoprotein. Also, there were clones identical to those selected as a protein associated with a disease.

[0061] The 243 clones in Table 21 are predicted to encode a glycoprotein. Among them, 53 clones were identical to those selected by the score in the ATGpr and the prediction by the PSORT for the existence of a signal sequence. And, many clones were turned out to be identical to the clones selected as a secretory or membrane protein. Moreover, there were clones identical to those selected as a protein associated with a disease.

[0062] The 51 clones in Table 22 are predicted to encode a protein associated to signal transduction.

[0063] The 130 clones in Table 23 are predicted to encode a protein associated to transcription.

[0064] The 17 clones in Table 24 are predicted to encode a protein associated with diseases.

[0065] In these clones, 532 clones have the maximal ATGpr1 score of 0.5 or higher (Table 25). 60 clones have the maximal ATGpr1 score of 0.3 or higher and less than 0.5 (Table 26 and NT2RP2006580). And 67 clones were with the maximal ATGpr1score of 0 or higher and less than 0.3 (Table 27).

[0066] 532 clones shown in Table 25, each having the maximal score in the ATGpr1 0.5 or higher, are very likely to be full-length and encode a secretory or membrane protein, or proteins associated to signal transduction, glycoprotein, transcription, or diseases. 59 clones in Table 26 and NT2RP2006580, which have the maximal score in the ATGpr1 0.3 or higher and less than 0.5, are likely to be full-length and encode a secretory or membrane protein, or proteins associated to signal transduction, glycoprotein, transcription, or diseases. 67 clones in Table 27, having the maximal score in the ATGpr1 0 or higher and less than 0.3, are still likely to be full-length and encode a secretory or membrane protein, or proteins associated to signal transduction, glycoprotein, transcription, or diseases.

[0067] This is the method for selecting the cDNA clones predicted to encode secretory and/or transmembrane proteins, glycoprotein-associated proteins, signal transduction-associated proteins, transcription-associated proteins, or disease-associated proteins on the basis of the partial sequences (5' sequences).

[0068] The polynucleotide of the present invention encodes an amino acid sequence of a functional protein such as a secretory or membrane protein, or a protein associated to signal transduction, glycoprotein, transcription, or diseases. Since the protein has the complete amino acid sequence, it is possible to analyze its biological activity by expressing the protein as a recombinant protein using an appropriate expression system, or by raising and using an antibody which specifically recognizes it.

[0069] It is possible to analyze the biological activity of a secretory protein or a membrane protein, or proteins associated to signal transduction, glycoprotein, or transcription, based on the methods in "Gene Transcription" (Hames B. D., and Higgins S.J. edit, (1993)), "Glycobiology" (Fukuda M., and Kobata A. edit, (1993)), "Growth Factors" (McKay I., and Leigh I. edit, (1993)), "Extracellular Matrix" (Haralson M.A., and Hassell J.R. edit, (1995), "Transcription Factors" (Latchman D.S. edit, (1993)), "Signal Transduction" (Milligans G. edit, (1992)), featured in "The Practical Approach

Series" (IRL PRESS), or "Signal Transduction Protocols" (Kendall D.A., and Hill S.J. edit, (1995), "Glycoprotein Analysis in Biomedicine" (Hounsell E.F. edit, (1993)), featured in "Method in Molecular Biology" (Humana Press).

[0070] As to a protein associated with a disease, it is possible to perform a functional analysis as described above, but also possible to analyze correlation between the expression or the activity of the protein and a certain disease by using a specific antibody that recognizes the protein. Alternatively, it is possible to utilize the database Online Mendelian Inheritance in Man (OMIM) (http://www.ncbi.nlm.nih.gov/Omim/), which is a database of human genes and diseases, to analyze the protein. New information is constantly being deposited in the OMIM database. Therefore, it is possible for one skilled in the art to find a new relationship between a particular disease and a gene of the present invention in the updated database.

[0071] Proteins associated with diseases are useful in drug development as they can be utilized as a diagnostic marker, a drug that regulates the level of their expression and activities, or a target of gene therapy. Also, as for a secretory protein, membrane protein, or proteins associated with signal transduction, glycoprotein, or transcription, search of the OMIM with the keywords mentioned below revealed that the proteins are associated with many diseases. Also, relationship between a proteins associated to signal transduction or transcription and diseases is reported in "Transcription Factor Research-1999" (Fujii, Tamura, Morohashi, Kageyama, and Satake edit, (1999) Jikken-Igaku Zoukan, Vol.17, No.3), and "Gene Medicine" ((1999) Vol.3, No.2). Thus, not only a protein associated with diseases, but also a secretory protein, membrane protein, or protein associated with signal transduction, glycoprotein, or transcription is involved in diseases, suggesting these proteins also are very important as a target in medical industry.

[0072] Keywords used in the search of the OMIM

(1) secretion protein
(2) membrane protein
(3) channel
(4) extracellular matrix
(5) receptor
(6) glycoprotein
(7) protein kinase
(8) calmodulin kinase
(9) transcription factor

[0073] Shown in the search result are only the accession numbers in the OMIM. Using the number, data showing the relationship between a disease and a gene or protein can be seen. The OMIM data has been renewed everyday.

1) Secretion protein
268 entries found, searching for "secretion protein"
104760, 176860, 160900, 107400, 118910, 139320, 603850, 147572, 176880, 600946, 603215, 157147, 600174, 151675, 170280, 179512, 179513, 138120, 179509, 246700, 179510, 600626, 179511, 600998, 109270, 601489, 154545, 179490, 185860, 603216, 122559, 601746, 147290, 602672, 146770, 603062, 179508, 131230, 601591, 602421, 139250, 167805, 167770, 600041, 600564, 118825, 601146, 300090, 600753, 601652, 600759, 600768, 602434, 182590, 603166, 308230, 602534, 603489, 107470, 150390, 104610, 173120, 158106, 143890, 306900, 308700, 134797, 137350, 227500, 176300, 107730, 600760, 138079, 120180, 120160, 120150, 124092, 138160, 101000, 227600, 600509, 601199, 142410, 104311, 193400, 201910, 107300, 122560, 272800, 217000, 590050, 147670, 133170, 176730, 300300, 134370, 274600, 120140, 162151, 158070,
152790, 120120, 106100, 300200, 192340, 190160, 138040, 147470, 147620, 173350, 147380, 152200, 152760, 157145, 153450, 264080, 113811, 600937, 600840, 188545, 202110, 600514, 186590, 603372, 136435, 137241, 252800, 214500, 207750, 138850, 139191, 142640, 138130, 189907, 603692, 600633, 603355, 107270, 600377, 147892, 232200, 600281, 232800, 602358, 137035, 601771, 601769, 253200, 601933, 118444, 600270, 120700, 600945, 603732, 147660, 600761, 172400, 600823, 600877, 130080, 171060, 107740, 307800, 602843, 130660, 152780, 124020, 601124, 601340, 601604, 601610, 171050, 312060, 232700, 300159, 142703, 600734, 125255, 168450, 123812, 188540, 147940, 188450, 600839, 182452, 188400, 182280, 176760, 263200, 600264, 188826, 252650, 601185, 162641, 137216, 601398, 601538, 118888, 118445, 601745,
190180, 601922, 182098, 602008, 147440, 602384, 600031, 109160, 602663, 151670, 602682, 602730, 602779, 146880, 603061, 142704, 603140, 106150, 600732, 153620, 603318, 139392, 600042, 102200, 603493, 182100, 264300, 603795, 184600

2) Membrane protein
1017 entries found, searching for "membrane protein"
130500, 305360, 153330, 173610, 170995, 109270, 170993, 309060, 120920, 602333, 133740, 133710, 602690,

603431, 147730, 603366, 603348, 600556, 602136, 164160, 310200, 152390,
601241, 116897, 137295, 600576, 194070, 601487, 600698, 164810, 601769, 141900, 602225, 275350, 131100,
179755, 600075, 162200, 165160, 116806, 600899, 123810, 133450, 216400, 278700, 190080, 164730, 191170,
193300, 600618, 600999, 601090, 106150, 601843, 133530, 110700, 602550, 138040, 133430, 300133, 163731,
602302, 126337, 309548, 180245, 126110, 602291, 109565, 107400, 314670, 601444, 143100, 104760, 106180,
601953, 600584, 125852, 602419, 600401, 142200, 107680, 167414, 600020, 188400, 208900, 175100, 602700,
601828, 139320, 602777, 600185, 602681, 603023, 314997, 602848, 600284, 102578, 114290, 165095, 137070,
602991, 602421, 600005, 602996, 314995, 152200, 151900, 112260, 129010, 600892, 273800, 176760, 602341,
490000, 136533, 400003, 601007, 602229, 603620, 602218, 602116, 602020,
142000, 601955, 126340, 120150, 193067, 182452, 142461, 194558, 180660, 600756, 160745, 107741, 106210,
157640, 186770, 146738, 603759, 213700, 147880, 152391, 277900, 151626, 107730, 600711, 600246, 107470,
600237, 223100, 107748, 600065, 600349, 426000, 194500, 154030, 227650, 600247, 314980, 109560, 305370,
600800, 301000, 277700, 600838, 312173, 600439, 600440, 191315, 601595, 190450, 190070, 190020, 162360,
131320, 133540, 600993, 601993, 159530, 601902, 602868, 181590, 601724, 602260, 601093, 187270, 164761,
602102, 603245, 136950, 106100, 601182, 167410, 601897, 602896, 170100, 602506, 104150, 176730, 601600,
187011, 102600, 180380, 162080, 603450, 142967, 602301, 126375, 603372, 603355, 164720, 603250, 167409,
167415, 602897, 601565, 185250, 182138, 601851, 600749, 601575, 194548,
154500, 601365, 194541, 601621, 601623, 601531, 600790, 194355, 123830, 123812, 154540, 601415, 143055,
601386, 194550, 186930, 131290, 601320, 601620, 601754, 601313, 184430, 182900, 182500, 600725, 147870,
154365, 116953, 601297, 601296, 601265, 600796, 120436, 601644, 601930, 601643, 230200, 601645, 601972,
600861, 602009, 601172, 601158, 601646, 180630, 600821, 118440, 601656, 601647, 150200, 601125, 601671,
141850, 116899, 600697, 109270, 202110, 150570, 601108, 191339, 601063, 109691, 180240, 203100, 151430,
179710, 111000, 176797, 238600, 104311, 240300, 125255, 600423, 158070, 602439, 600324, 112261, 243305,
602474, 174762, 600613, 602539, 138890, 138720, 114550, 173865, 602582, 602584, 173510, 600250, 602627,
173325, 602635, 246530, 172425, 600193, 602691, 600188, 170998, 152790,
168468, 256540, 225250, 600848, 143400, 168461, 262600, 168360, 601912, 602951, 600017, 230000, 266600,
602981, 272800, 109150, 102200, 603025, 603026, 603109, 167050, 603127, 603128, 165240, 230400, 313700,
164975, 164875, 602017, 115500, 235800, 164873, 602110, 164785, 164772, 312865, 603296, 600542, 164740,
602125, 309801, 602148, 300007, 306955, 603368, 116940, 602181, 603416, 126650, 163920, 300024, 603437,
602209, 603576, 603607, 305435, 600944, 180410, 303630, 159557, 301870, 132810, 100790, 603849, 603862,
603881,

[0074] There are several methods for analyzing the expression levels of genes associated with diseases. Differences in gene expression levels between diseased and normal tissues are studied by the analytical methods, for example, Northern hybridization and differential display. Other examples include a method with high-density cDNA filter, a method with DNA microarray and methods with PCR amplification (Experimental Medicine, Vol.17, No. 8, 980-1056 (1999); Cell Engineering (additional volume) DNA Microarray and Advanced PCR Methods, Muramatsu & Naba (eds.), Shujunsya). The levels of gene expression between diseased tissues and normal tissues can be studied by any of these analytical methods. When explicit difference in expression level is observed for a gene, it can be concluded that the gene is closely associated with a disease or disorder. Instead of diseased tissues, cultured cells can be used for the assessment. Similarly, when gene expression is explicitly different between normal cells and cells reproducing disease-associated specific features, it can be concluded that the gene is closely associated with a disease or disorder. When the expression levels of genes are evidently varied during major cellular events (such as differentiation and apoptosis), the genes are involved in the cellular events and accordingly are candidates for disease- and/or disorder-associated genes. Further, genes exhibiting tissue-specific expression are genes playing important parts in the tissue functions and, therefore, can be candidates for genes associated with diseases and/or disorders affecting the tissues.

[0075] For example, non-enzymic protein glycation reaction is believed to be a cause for a variety of chronic diabetic complications. Accordingly, in endothelial cells, genes, of which expression levels are elevated or decreased in a glycated protein-dependent manner, are associated with diabetic complications caused by glycated proteins (Diabetes 1996, 45 (Suppl. 3), S67-S72; Diabetes 1997, 46 (Suppl. 2), S19-S25). The onset of rheumatoid arthritis is thought to be involved in the proliferation of synovial cells covering inner surfaces of joint cavity and in inflammatory reaction resulted from the action of cytokines produced by leukocytes infiltrating into the joint synovial tissues (Rheumatism Information Center, http://www.rheuma-net.or.jp/). Recent studies have also revealed that tissue necrosis factor (TNF)-α participates in the onset (Current opinion in immunology 1959, 11, 657-662). When the expression of a gene exhibits responsiveness to the action of TNF on synovial cells, the gene is considered to be involved in rheumatoid arthritis. Many genes acting at the downstream of TNF-. and IL-1. among inflammation-associated cytokines have been previously identified. The respective stimulations are transduced through independent pathways of signaling cascade. There exists another signaling cascade for both stimulations, wherein NF-.B is a common transducing molecule shared by

the two stimulations (J. Leukoc. Biol., 1994, 56(5): 542-547). It has also been revealed that many inflammation-associated genes, including IL-2, IL-6 and G-CSF, are varied in the expression levels in response to the signal through the common pathway (Trends Genet. 1999, 15(6): 229-235). It is assumed that genes of which expression levels are varied in response to the stimulation of TNF-. or IL-1. also participate in inflammation. Genes associated with neural differentiation can be candidates for causative genes for neurological diseases as well as candidates for genes usable for treating the diseases.

[0076] Clones exhibiting differences in the expression levels thereof can be selected by using gene expression analysis. The selection comprises, for example; analyzing cDNA clones by using high-density cDNA filter; and statistically treating the multiple signal values (signal values of radioisotope in the radiolabeled probes or values obtained by measuring fluorescence intensities emitted from the fluorescent labels) for the respective clones by two-sample t-test, where the signal values are determined by multiple experiments of hybridization. The clones of interest are selectable based on the statistically significant differences in the signal distribution at $p<0.05$. However, selectable clones with significant difference in the expression levels thereof may be changed depending on the partial modification of statistical treatment. For example, the clones may be selected by conducting statistical treatment with two-sample t-test at $p<0.01$; or genes exhibiting more explicit differences in the expression levels thereof can be selected by performing statistical treatment with a pre-determined cut-off value for the significant signal difference. An alternative method is that the expression levels are simply compared with each other, and then, the clones of interest are selected based on the ratio of the expression levels thereof.

[0077] Clones exhibiting differences in the expression levels can also be selected by comparing the expression levels by PCR analysis, for example, by using the method of determining the band intensities representing the amounts of PCR products with ethidium bromide staining; the method of determining the radioisotopic signal values or fluorescence intensities of the PCR products when radio-labeled or fluorescence-labeled primers; or the method of determining the values of radioisotope signals or fluorescence intensities of the probes hybridized to the PCR products when radio-labeled or fluorescence-labeled probes, respectively, are used in the hybridization. If the expression level ratios obtained in multiple PCR experiments are constantly at least 2-fold, such a clone can be judged to exhibit the difference in the expression level. When the ratios are several-fold or not less than 10-fold, the clone can be selected as a gene exhibiting the explicit difference in the expression level.

[0078] A survey of genes of which expression levels are varied specifically to the glycated protein in the endothelial cells revealed three genes with elevated expression levels, NT2RP2001538, NT2RP4001001 and Y79AA1000967. These clones are genes associated with diabetes.

[0079] A survey of genes of which expression levels are varied in response to TNF. (Tumor Necrosis Factor-alpha) in the primary cell culture of synovial tissue detected the following clones with elevated expression levels in the presence of TNF.:

BNGH41000020, HEMBA1000349, HEMBA1000634, HEMBA1000671, HEMBA1000835, HEMBA1000962, HEMBA1002178, HEMBA1002195, HEMBA1002239, HEMBA1002420, HEMBA1002524, HEMBA1002992, HEMBA1003315, HEMBA1003392, HEMBA1003487, HEMBA1003602, HEMBA1004067, HEMBA1004797, HEMBA1005337, HEMBA1005489, HEM3A1006916, HEMBB1000668, HEMBB1000905, HEMBB1001547, HEMBB1001573, HEMBB1002041, HEMBB1002663, MAMMA1000652, MAMMA1000810, MAMMA1001634, MAMMA1002091, MAMMA1002234, NT2RM2000306, NT2RM4000417, NT2RP1000002, NT2RP1000181, NT2RP1000740, NT2RP2000694, NT2RP2001921, NT2RP2002527, NT2RP2004495, NT2RP2004606, NT2RP2005163, NT2RP2005463, NT2RP2006134, NT2RP3000171, NT2RP3000652, NT2RP3001195, NT2RP3001976, NT2RP3003473, NT2RP3003874, NT2RP3004090, NT2RP3004294, NT2RP3004557, NT2RP3004647, NT2RP4000108, NT2RP4001001, NT2RP4001877, OVARC1000090, OVARC1000105, OVARC1000275, OVARC1000439, OVARC1001607, PLACE1000740, PLACE1000927, PLACE1001016, PLACE1001100, PLACE1001464, PLACE1001500, PLACE1001918, PLACE1002095, PLACE1002547, PLACE1003644, PLACE1004519, PLACE1005031, PLACE1005410, PLACE1005736, PLACE1006219, PLACE1006809, PLACE1008716, PLACE1010081, THYRO1001770, Y79AA1000127, Y79M1000207, Y79AA1000270, Y79AA1000876, Y794A1001013, Y79AA1001264, Y79AA1001272, Y79AA1001328, Y79AA1001430, Y79AA1001530, Y79AA1001799

[0080] Clones with decreased expression levels in the presence of TNF□ are NT2RM4000326, NT2RP1000300, NT2RP2000514, NT2RP2001755, NT2RP2006042, NT2RP3000481, NT2RP3002790. These clones are candidates for rheumatoid arthritis-associated genes.

[0081] A survey of genes of which expression levels are varied in response to TNF. (Tumor Necrosis Factor-alpha) or IL-1. (Interleukin-1 beta) in a human T cell strain, Jurkat cell, revealed the following clones with elevated expression levels in the presence of TNF.:

MAMMA1000141, MAMMA1000788, MAMMA1001237, MAMMA1002070,NT2RM2000582, NT2RM2002109, NT2RP1000679, NT2RP2003664, NT2RP2004108, NT2RP2005597, NT2RP3001592, NT2RP3002738, NT2RP3004133, NT2RP3004294, NT2RP3004321, NT2RP3004557, PLACE1002547, PLACE1003573,

PLACE1004305, PLACE1008744, PLACE1010011, PLACE1010713, PLACE1011181, Y79AA1000776, Y79AA1002129

[0082] The survey also revealed a clone of which expression level was decreased in the presence of TNF. The clone is PLACE1002070. The same survey further revealed the clones of which expression levels were elevated in the presence of IL-1.. The clones are MAMMA1000614, MAMMA1001237, NT2RM2000514 and NT2RP3001159. These clones are genes associated with inflammation.

[0083] A survey of genes of which expression levels are varied in response to the stimulation for inducing cell differentiation (stimulation using retinoic acid (RA) or using RA/inhibitor (inhibitor for cell division)) in cultured cells of neural strain, NT2, revealed the following clones with elevated expression levels in the presence of RA:

HEMBA1000121, HEMBA1000275, HEMBA1000300, HEMBA1000634, HEMBA1000671, HEMBA1000875, HEMBA1001184, HEMBA1001390, HEMBA1001886, HEMBA1002163, HEMBA1002227, HEMBA1002420, HEMBA1002421, HEMBA1003072, HEMBA1003120, HEMBA1003294, HEMBA1003497, HEMBA1004007, HEMBA1004110, HEMBA1004391, HEMBA1004444, HEMBA1005230, HEMBA1005246, HEMBA1005267, HEMBA1005489, HEMBA1005913, HEMBA1006299, HEMBA1006357, HEMBA1006517, HEMBA1006544, HEMBA1006658, HEMBA1006749, HEMBA1007063, HEMBA1007241, HEMBB1000447, HEMBB1000542, HEMBB1000567, HEMBB1000642, HEMBB1000668, HEMBB1001026, HEMBB1001847, HEMBB1002051, HEMBB1002120, HEMBB1002228, HEMBB1002693, MAMMA1000106, MAMMA1000141, MAMMA1000473, MAMMA1000528, MAMMA1000810, MAMMA1000881, MAMMA1001634, MAMMA1001957, MAMMA1002205, MAMMA1002224, NT2RM2000423, NT2RM2000497, NT2RM2000582, NT2RM2001126, NT2RM2001902, NT2RM4000198, NT2RM4000284, NT2RM4000593, NT2RM4001321, NT2RP1000002, NT2RP1000050, NT2RP1000181, NT2RP1000261, NT2RP1000465, NT2RP1000468, NT2RP1000579, NT2RP1000679, NT2RP2000092, NT2RP2000479, NT2RP2000610, NT2RP2000663, NT2RP2000694, NT2RP2000903, NT2RP2001388, NT2RP2001538, NT2RP2001878, NT2RP2002015, NT2RP2002304, NT2RP2002721, NT2RP2002824, NT2RP2002942, NT2RP2002974, NT2RP2002976, NT2RP2003179, NT2RP2003302, NT2RP2003383, NT2RP2003469, NT2RP2003664, NT2RP2003940, NT2RP2004069, NT2RP2004108, NT2RP2004524, NT2RP2004556, NT2RP2004670, NT2RP2005069, NT2RP2005247, NT2RP2005425, NT2RP2005463, NT2RP2005514, NT2RP2005535, NT2RP2005541, NT2RP2005774, NT2RP2005878, NT2RP2005883, NT2RP2005887, NT2RP2006099, NT2RP2006134, NT2RP3000011, NT2RP3000125, NT2RP3000171, NT2RP3000232, NT2RP3000460, NT2RP3000481, NT2RP3000652, NT2RP3000677, NT2RP3000818, NT2RP3000820, NT2RP3001044, NT2RP3001061, NT2RP3001170, NT2RP3001240, NT2RP3001322, NT2RP3001388, NT2RP3001542, NT2RP3001592, NT2RP3001976, NT2RP3002790, NT2RP3002900, NT2RP3002983, NT2RP3003000, NT2RP3003354, NT2RP3003532, NT2RP3003729, NT2RP3003874, NT2RP3003939, NT2RP3004025, NT2RP3004083, NT2RP3004090, NT2RP3004130, NT2RP3004202, NT2RP3004294, NT2RP3004640, NT2RP4000108, NT2RP4000634, NT2RP4002451, NT2RP4002715, OVARC1000090, OVARC1000208, OVARC1000275, OVARC1000553, OVARC1000775, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000995, OVARC1001030, OVARC1001049, OVARC1001132, OVARC1001596, OVARC1002178, PLACE1000258, PLACE1000442, PLACE1000927, PLACE1000986, PLACE1001100, PLACE1001123, PLACE1001795, PLACE1002518, PLACE1002547, PLACE1002967, PLACE1003407, PLACE1003428, PLACE1003644, PLACE1003839, PLACE1004078, PLACE1004441, PLACE1004450, PLACE1005669, PLACE1005682, PLACE1005736, PLACE1005768, PLACE1005815, PLACE1006073, PLACE1006208, PLACE1007296, PLACE1007626, PLACE1008282, PLACE1008984, PLACE1008985, PLACE1010445, PLACE1011708, PLACE1011978, PLACE4000455, SKNMC1000004, THYRO1000036, THYRO1000580, THYRO1000776, THYRO1000999, THYRO1001063, THYRO1001128, THYRO1001205, THYRO1001327, THYRO1001523, THYRO1001725, THYRO1001770, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1001056, Y79AA1001062, Y79AA1001090, Y79AA1001727, Y79AA1002213, Y79AA1002381

[0084] The survey also revealed the clones of which expression levels were decreased in the presence of RA. The clones are BNGH41000020, HEMBA1005070, NT2RP2005027, NT2RP3003473 and Y79AA1002376.

[0085] The same survey further revealed the following clones with elevated expression levels in the presence of RA/inhibitor:

HEMBA1000128, HEMBA1000875, HEMBA1001390, HEMBA1002163, HEMBA1002227, HEMBA1002421, HEMBA1004391, HEMBA1004454, HEMBA1004785, HEMBA1005913, HEMBA1006171, HEMBA1006299, HEMBA1006335, HEMBA1006544, HEMBA1007241, HEMBB1000447, HEMBB1000668, MAMMA1000994, MAMMA1001344, NT2RM2000582, NT2RP1001004, NT2RP2000663, NT2RP2000694, NT2RP2000903, NT2RP2001388, NT2RP2002674, NT2RP2002974, NT2RP2003383, NT2RP2004069, NT2RP2004606, NT2RP2004837, NT2RP2005069, NT2RP2005425, NT2RP2005463, NT2RP2005541, NT2RP2005883, NT2RP2005887, NT2RP3000460, NT2RP3000838, NT2RP3001044, NT2RP3001240, NT2RP3001388,

NT2RP3002721, NT2RP3002738, NT2RP3003469, NT2RP3004083, NT2RP3004130, NT2RP3004202, NT2RP3004294, NT2RP3004640, NT2RP4000108, NT2RP4002451, NT2RP4002715, OVARC1000275, OVARC1000467, OVARC1000553, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000995, OVARC1001030, OVARC1001222, OVARC1001596, OVARC1002058, OVARC1002178, PLACE1000927, PLACE1001123, PLACE1001407, PLACE1001464, PLACE1001564, PLACE1001795, PLACE1002547, PLACE1003407, PLACE1003644, PLACE1003845, PLACE1004441, PLACE1004482, PLACE1005410, PLACE1005601, PLACE1005725, PLACE1005736, PLACE1006093, PLACE1006219, PLACE1006290, PLACE1006716, PLACE1007296, PLACE1007626, PLACE1008359, PLACE1010968, PLACE1011364, PLACE1011824, THYRO1000678, THYRO1000776, THYRO1000999, THYRO1001113, THYRO1001237, THYRO1001523, Y79AA1000226, Y79AA1000888, Y79AA1001430

[0086] The same survey further revealed the following clones with elevated expression levels in the presence of RA/inhibitor:

HEMBA1000349, HEMBA1001297, HEMBA1001878, HEMBA1005070, HEMBA1006482, HEMBB1001959, NT2RM2001939, NT2RP1000981, NT2RP2001469, NT2RP3003473, OVARC1001132, PLACE1001655, Y79AA1000127, Y79AA1002381. These clones are associated with neural differentiation and, therefore, are candidates for genes associated with neurological diseases.

[0087] Based on the functional analyses using a secretory protein, membrane protein, or proteins associated with signal transduction, glycoprotein, transcription, or diseases, it is possible to develop a medicine.

[0088] In case of a membrane protein, it is most likely to be a protein that functions as a receptor or ligand on the cell surface. Therefore, it is possible to reveal a new relationship between a ligand and receptor by screening the membrane protein of the invention based on the binding activity with the known ligand or receptor. Screening can be performed according to the known methods.

[0089] For example, a ligand against the protein of the invention can be screened in the following manner. Namely, a ligand that binds to a specific protein can be screened by a method comprising the steps of: (a) contacting a test sample with the protein of the invention or a partial peptide thereof, or cells expressing these, and (b) selecting a test sample that binds to said protein, said partial peptide, or said cells.

[0090] On the other hand, for example, screening using cells expressing the protein of the present invention that is a receptor protein can also be performed as follows. It is possible to screen receptors that is capable of binding to a specific protein by using procedures (a) attaching the sample cells to the protein of the invention or its partial peptide, and (b) selecting cells that can bind to the said protein or its partial peptide.

[0091] In a following screening as an example, first the protein of the invention is expressed, and the recombinant protein is purified. Next, the purified protein is labeled, binding assay is performed using a various cell lines or primary cultured cells, and cells that are expressing a receptor are selected (Growth and differentiation factors and their receptors, Shin-Seikagaku Jikken Kouza Vol.7 (1991) Honjyo, Arai, Taniguchi, and Muramatsu edit, p203-236, Tokyo-Kagaku-Doujin). A protein of the invention can be labeled with RI such as ^{125}I, and enzyme (alkaline phosphatase etc.). Alternatively, a protein of the invention may be used without labeling and then detected by using a labeled antibody against the protein. The cells that are selected by the above screening methods, which express a receptor of the protein of the invention, can be used for the further screening of an agonists or antagonists of the said receptor.

[0092] Once the ligand binding to the protein of the invention, the receptor of the protein of the invention or the cells expressing the receptor are obtained by screening, it is possible to screen a compound that binds to the ligand and receptor. Also it is possible to screen a compound that can inhibit both bindings (agonists or antagonists of the receptor, for example) by utilizing the binding activities.

[0093] When the protein of the invention is a receptor, the screening method comprises the steps of (a) contacting the protein of the invention or cells expressing the protein of the invention with the ligand, in the presence of a test sample, (b) detecting the binding activity between said protein or cells expressing said protein and the ligand, and (c) selecting a compound that reduces said binding activity when compared to the activity in the absence of the test sample. Furthermore, when the protein of the invention is a ligand, the screening method comprises the steps of (a) contacting the protein of the invention with its receptor or cells expressing the receptor in the presence of samples, (b) detecting the binding activity between the protein and its receptor or the cells expressing the receptor, and (c) selecting a compound that can potentially reduce the binding activity compared to the activity in the absence of the sample.

[0094] Samples to screen include cell extracts, expressed products from a gene library, synthesized low molecular compound, synthesized peptide, and natural compounds, for example, but are not construed to be listed here. A compound that is isolated by the above screening using a binding activity of the protein of the invention can also be used as a sample.

[0095] A compound isolated by the screening may be a candidate to be an agonist or an antagonist of the receptor of the protein. By utilizing an assay that monitors a change in the intracellular signaling such as phosphorylation which results from reduction of the binding between the protein and its receptor, it is possible to identify whether the obtained compound is an agonist or antagonist of the receptor. Also, the compound may be a candidate of a molecule that can

inhibit the interaction between the protein and its associated proteins (including a receptor) in vivo. Such compounds can be used for developing drugs for precaution or cures of a disease with which the protein is associated.

[0096] Secretory proteins may regulate cellular conditions such as growth and differentiation. It is possible to find out a novel factor that regulates cellular conditions by adding the secretory protein of the invention to a certain kind of cell, and performing a screening by utilizing the cellular changes in growth or differentiation, or activation of a particular gene.

[0097] The screening can be performed, for example, as follows. First, the protein of the invention is expressed and purified in a recombinant form. Then, the purified protein is microinjectedadded into a various kind of cell lines or primary cultured cells, and the change in the cell growth and differentiation is monitored. The induction of a particular gene that is known to be involved in a certain cellular change is detected with the amounts of mRNA and protein. Alternatively, the amount of an intracellular molecule (low molecular compounds, etc.) that is changed by the function of a gene product (protein) that is known to be functioning in a certain cellular change is used for the detection.

[0098] Once the screening reveals that the protein of the invention can regulate cellular conditions or the functions, it is possible to apply the protein as a pharmaceutical and diagnostic medicine for associated diseases by itself or by altering a part of it into an appropriate composition.

[0099] As is above described for membrane proteins, the secretory protein provided by the invention may be used to explore a novel ligand-receptor interaction using a screening based on the binding activity to a known ligand or receptor. A similar method can be used to identify an agonist or antagonist. The resulting compounds obtained by the methods can be a candidate of a compound that can inhibit the interaction between the protein of the invention and an interacting molecule (including a receptor). The compounds may be able to use as a preventive, therapeutic, and diagnostic medicine for the diseases, in which the protein may play a certain role.

[0100] Proteins associated with signal transduction or transcription may be a factor that affects a certain protein or gene in response to intracellular/extracellular stimuli. It is possible to find out a novel factor that can affect a protein or gene by expressing the protein provided by the invention in a certain types of cells, and performing a screening utilizing the activation of a certain intracellular protein or gene.

[0101] The screening may be performed as follows. First, a transformed cell expressing the protein is obtained. Then, the transformed cell line and the untransformed original cell are compared for the changes in the expression of a certain gene by detecting the amount of its mRNA or protein. Alternatively, the amount of an intracellular molecule (low molecular compounds), which is changed by the function of a gene product (protein) that is known to function in a certain cellular change, may be used for the detection. Furthermore, the change of the expression of a certain gene can be detected by introducing a fusion gene that comprises a regulatory region of the gene and a marker gene (luciferase, beta-galactosidase, etc.) into a cell, expressing the protein provided by the invention into the cell, and estimating the activity of a marker gene product (protein).

[0102] If the protein or gene of the invention is associated with diseases, it is possible to screen a gene or compound that can regulate its expression and/or activity either directly or indirectly by utilizing the protein of the present invention.

[0103] For example, the protein of the invention is expressed and the recombinant protein is purified. Then, the protein and gene whose expression was affected in the presence of the protein of the invention is also purified, and the binding activity between the two proteins or genes is examined. The examination may be performed with pretreatment with a compound that is candidate of an inhibitor. In an alternative method, a transcription regulatory region locating in the 5'-upstream of the gene encoding the protein of the invention that is capable of regulating the expression of other genes is obtained, and fused with a marker gene. The fusion is introduced into a cell, and the cell is added with compounds to explore a regulatory factor of the expression of the said gene.

[0104] The compound obtained by the screening can be used for developing pharmaceutical and diagnostic medicines for the diseases with which the protein of the present invention is associated. Similarly, if the regulatory factor obtained by the screening is a protein, the protein itself can be used as a pharmaceutical, and if there is a compound that affects the original expression level and/or activity of the protein, it also can be used for the same purpose.

[0105] If the protein of the invention has an enzymatic activity, regardless of whether it is a secretory protein, membrane protein, or proteins associated with signal transduction, glycoprotein, transcription, or diseases, a screening may be performed by adding a compound to the protein of the invention under the suitable condition and monitoring the change of the compound. The enzymatic activity may also be utilized to screen for a compound that can inhibit the activity of the protein.

[0106] In a screening given as an example, the protein of the invention is expressed and the recombinant protein is purified. Then, compounds are contacted with the purified protein, and the amount of the compound and the reaction products is examined. Alternatively, compounds that are candidates of an inhibitor are pretreated, then a compound (substrate) that can react with the purified protein is added, and the amount of the substrate and the reaction products is examined.

[0107] The compounds obtained in the screening may be used as a medicine for diseases with which the protein of the invention is associated. Also they can be applied for tests that examine whether the protein of the invention functions

normally *in vivo*.

[0108] Whether the secretory or membrane protein of the present invention is a novel protein associated with diseases or not is determined in another method than described above, by obtaining a specific antibody against the protein of the invention, and examining the relationship between the expression or activity of the protein and a certain disease. In an alternative way, it may be analyzed referred to the methods in "Molecular Diagnosis of Genetic Diseases" (Elles R. edit, (1996) in the series of "Method in Molecular Biology" (Humana Press).

[0109] Disease-associated proteins are a target of the above described screenings and very useful for developing a drug that is capable of regulating the expression and activity of the protein. Also, they are useful in medicinal industry as a diagnostic marker of the related disease and as a target for gene therapy.

[0110] Compounds isolated as mentioned above can be administered patients as it is, or after formulated into a pharmaceutical composition according to the known methods. For example, a pharmaceutically acceptable carrier or vehicle, specifically sterilized water, saline, plant oil, emulsifier, or suspending agent can be mixed with the compounds appropriately. The pharmaceutical compositions can be administered to patients by a method known to those skilled in the art, such as intraarterial intravenous, or subcutaneous injections. The dosage may vary depending on the weight or age of a patient, or the method of administration, but those skilled in the art can choose an appropriate dosage properly. If the compound is encoded by DNA, the DNA can be cloned into a vector for gene therapy, and used for gene therapy. The dosage of the DNA and the method of its administration may vary depending on the weight or age of a patient, or the symptoms, but those skilled in the art can choose properly.

[0111] The protein encoded by the polynucleotide of the invention can be prepared as a recombinant protein or as a natural protein. For example, the recombinant protein can be prepared by inserting the polynucleotide encoding the protein of the invention into a vector, introducing the vector into an appropriate host cell and purifying the protein expressed within the transformed host cell, as described below. In contrast, the natural protein can be prepared, for example, by utilizing an affinity column to which an antibody against the protein of the invention (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wily & Sons, Section 16.1-16.19) is attached. The antibody used for the affinity purification may be either a polyclonal antibody, or a monoclonal antibody. Alternatively, in vitro translation (See, for example, "On the fidelity of mRNA translation in the nuclease-treated rabbit reticulocyte lysate system." Dasso M.C., and Jackson R.J. (1989) Nucleic Acids Res. 17: 3129-3144) may be used for preparing the protein of the invention.

[0112] Proteins functionally equivalent to the proteins of the present invention can be prepared based on the activities, which were clarified in the above-mentioned manner, of the proteins of the present invention. Using the biological activity possessed by the protein of the invention as an index, it is possible to verify whether or not a particular protein is functionally equivalent to the protein of the invention by examining whether or not the protein has said activity.

[0113] Proteins functionally equivalent to the proteins of the present invention can be prepared by those skilled in the art, for example, by using a method for introducing mutations into an amino acid sequence of a protein (for example, site-directed mutagenesis (Current Protocols in Molecular Biology, edit, Ausubel et al., (1987) John Wiley & Sons, Section 8.1-8.5). Besides, such proteins can be generated by spontaneous mutations. The present invention comprises the proteins having one or more amino acid substitutions, deletions, insertions and/or additions in the amino acid sequences of the proteins of the present invention (Table 370), as far as the proteins have the equivalent functions to those of the proteins identified in the present Examples described later.

[0114] There are no limitations in the number and sites of amino acid mutations, as far as the proteins maintain the functions thereof. The number of mutations is typically 30% or less, or 20% or less, or 10% or less, preferably within 5% or less, or 3% or less of the total amino acids, more preferably within 2% or less or 1% or less of the total amino acids. From the viewpoint of maintaining the protein function, it is preferable that a substituted amino has a similar property to that of the original amino acid. For example, Ala, Val, Leu, Ile, Pro, Met, Phe and Trp are assumed to have similar properties to one another because they are all classified into a group of non-polar amino acids. Similarly, substitution can be performed among non-charged amino acid such as Gly, Ser, Thr, Cys, Tyr, Asn, and Gln, acidic amino acids such as Asp and Glu, and basic amino acids such as Lys, Arg, and His.

[0115] In addition, proteins functionally equivalent to the proteins of the present invention can be isolated by using techniques of hybridization or gene amplification known to those skilled in the art. Specifically, using the hybridization technique (Current Protocols in Molecular Biology, edit, Ausubel et al., (1987) John Wiley & Sons, Section 6.3-6.4)), those skilled in the art can usually isolate a DNA highly homologous to the DNA encoding the protein identified in the present Example based on the identified nucleotide sequence (Table 370) or a portion thereof and obtain the functionally equivalent protein from the isolated DNA. The present invention includes proteins encoded by the DNAs hybridizing with the DNAs encoding the proteins identified in the present Example, as far as the proteins are functionally equivalent to the proteins identified in the present Example. Organisms from which the functionally equivalent proteins are isolated are illustrated by vertebrates such as human, mouse, rat, rabbit, pig and bovine, but are not limited to these animals.

[0116] Washing conditions of hybridization for the isolation of DNAs encoding the functionally equivalent proteins are usually "1xSSC, 0.1% SDS, 37"; more stringent conditions are "0.5xSSC, 0.1% SDS, 42."; and still more stringent conditions are "0.1 x SSC, 0.1% SDS, 65.". Alternatively, the following conditions can be given as hybridization con-

ditions of the present invention. Namely, conditions in which the hybridization is done at "6xSSC, 40% Formamide, 25.", and the washing at "1xSSC, 55." can be given. More preferable conditions are those in which the hybridization is done at "6xSSC, 40% Formamide, 37.", and the washing at "0.2xSSC, 55,". Even more preferable are those in which the hybridization is done at "6xSSC, 50% Formamide, 37.", and the washing at "0.1xSSC, 62.". The more stringent the conditions of hybridization are, the more frequently the DNAs highly homologous to the probe sequence are isolated. Therefore, it is preferable to conduct hybridization under stringent conditions. Examples of stringent conditions in the present invention are, washing conditions of "0.5xSSC, 0.1% SDS, 42.", or alternatively, hybridization conditions of "6xSSC, 40% Formamide, 37.", and the washing at "0.2xSSC, 55.". However, the above-mentioned combinations of SSC, SDS and temperature conditions are indicated just as examples. Those skilled in the art can select the hybridization conditions with similar stringency to those mentioned above by properly combining the above-mentioned or other factors (for example, probe concentration, probe length and duration of hybridization reaction) that determines the stringency of hybridization.

[0117] The amino acid sequences of proteins isolated by using the hybridization techniques usually exhibit high homology to those of the proteins of the present invention, which are shown in Table 370. The present invention encompasses a polynucleotide comprising a nucleotide sequence that has a high identity to the nucleotide sequence of claim 8 (a). Furthermore, the present invention encompasses a peptide, or protein comprising an amino acid sequence that has a high identity to the amino acid sequence encoded by the polynucleotide of claim 8 (b). The term "high identity" indicates sequence identity of at least 40% or more; preferably 60% or more; and more preferably 70% or more. Alternatively, more preferable is identity of 90% or more, or 93% or more, or 95% or more, furthermore, 97% or more, or 99% or more. The identity can be determined by using the BLAST search algorithm.

[0118] With the gene amplification technique (PCR) (Current Protocols in Molecular Biology, edit, Ausubel et al., (1987) John Wiley & Sons, Section 6.3-6.4)) using primers designed based on the nucleotide sequence (Table 370) or a portion thereof identified in the present Example, it is possible to isolate a DNA fragment highly homologous to the nucleotide sequence or a portion thereof and to obtain functionally equivalent protein to a particular protein identified in the present Example based on the isolated DNA fragment.

[0119] The "percent identity" of two amino acid sequences or of two nucleic acids is determined using the algorithm of Karlin and Altschul (Proc. Natl. Acad. Sei. USA 87:2264-2268, 1990), modified as in Karlin and Altschul (Proc. Natl. Acad. Sei. USA 90:5873-5877, 1993). Such an algorithm is incorporated into the BLASTN and BLASTX programs of Altschul et al. (J. Mol. Biol.215:403-410, 1990). BLAST nucleotide searches are performed with the BLASTN program, score = 100, wordlength = 12. BLAST protein searches are performed with the BLASTX program, score = 50, wordlength = 3. When gaps exist between two sequences, Gapped BLAST is utilized as described in Altschul et al. (Nucleic Acids Res.25:3389-3402,1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) are used. See http://www.ncbi.nlm.nih.gov.

[0120] The present invention also includes a partial peptide of the proteins of the invention. The partial peptide comprises a protein generated as a result that a signal peptide has been removed from a secretory protein. If the protein of the present invention has an activity as a receptor or a ligand, the partial peptide may function as a competitive inhibitor of the protein and may bind to the receptor (or ligand). In addition, the present invention comprises an antigen peptide for raising antibodies. For the peptides to be specific for the protein of the invention, the peptides comprise at least 7 amino acids, preferably 8 amino acids or more, more preferably 9 amino acids or more, and even more preferably 10 amino acids or more. The peptide can be used for preparing antibodies against the protein of the invention, or competitive inhibitors of them, and also screening for a receptor that binds to the protein of the invention. The partial peptides of the invention can be produced, for example, by genetic engineering methods, known methods for synthesizing peptides, or digesting the protein of the invention with an appropriate peptidase.

[0121] The present invention also relates to a vector into which the DNA of the invention is inserted. The vector of the invention is not limited as long as it contains the inserted DNA stably. For example, if E. coli is used as a host, vectors such as pBluescript vector (Stratagene) are preferable as a cloning vector. To produce the protein of the invention, expression vectors are especially useful. Any expression vector can be used as far as it is capable of expressing the protein in vitro, in E. coli, in cultured cells, or in vivo. For example, pBEST vector (Promega) is preferable for in vitro expression, pET vector (Invitrogen) for E. coli, pME18S-FL3 vector (GenBank Accession No. AB009864) for cultured cells, and pME18S vector (Mol. Cell. Biol. (1988) 8: 466-472) for in vivo expression. To insert the DNA of the invention, ligation utilizing restriction sites can be performed according to the standard method (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wily & Sons, Section 11.4-11.11).

[0122] The present invention also relates to a transformant carrying the vector of the invention. Any cell can be used as a host into which the vector of the invention is inserted, and various kinds of host cells can be used depending on the purposes. For strong expression of the protein in eukaryotic cells, COS cells or CHO cells can be used, for example.

[0123] Introduction of the vector into host cells can be performed, for example, by calcium phosphate precipitation method, electroporation method (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wily & Sons, Section 9.1-9.9), lipofectamine method (GIBCO-BRL), or microinjection method, etc.

[0124] The primer of the present invention can be used for synthesizing full-length cDNA, and also for the detection and/or diagnosis of the abnormality of the protein of the invention encoded by the full-length cDNA. For example, by utilizing polymerase chain reaction (genomic DNA-PCR, or RT-PCR) using the primer of the invention, DNA encoding the protein of the invention can be amplified. It is also possible to obtain the regulatory region of expression in the 5'-upstream by using PCR or hybridization since the transcription start site within the genomic sequence can be easily specified based on the 5'-end sequence of the full-length cDNA. The obtained genomic region can be used for detection and/or diagnosis of the abnormality of the sequence by RFLP analysis, SSCP, or direct sequencing.

[0125] Furthermore, the "polynucleotide having a length of at least 15 nucleotides, comprising a nucleotide sequence that is complementary to a polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs in Table 370, or its complementary strand" includes an antisense polynucleotide for suppressing the expression of the protein of the invention. To exert the antisense effect, the antisense polynucleotide has a length of at least 15 bp or more, for example, 50 bp or more, preferably 100 bp or more, and more preferably 500 bp or more, and has a length of usually 3000 bp or less and preferably 2000 bp or less. The antisense DNA can be used in the gene therapy of the diseases that are caused by the abnormality of the protein of the invention (abnormal function or abnormal expression). Said antisense DNA can be prepared, for example, by the phosphorothioate method ("Physicochemical properties of phosphorothioate oligodeoxynucleotides." Stein (1988) Nucleic Acids Res. 16: 3209-3221) based on the nucleotide sequence of the DNA encoding the protein (for example, the DNA set forth in any one of SEQ ID NOs in Table 370).

[0126] The polynucleotide or antisense DNA of the present invention can be used in gene therapy, for example, by administrating it into a patient by the in vivo or ex vivo method with virus vectors such as retrovirus vectors, adenovirus vectors, and adeno-associated virus vectors, or non-virus vectors such as liposome.

[0127] The present invention also relates to antibodies that bind to the protein of the invention. There are no limitations in the form of the antibodies of the invention. They include polyclonal antibodies, monoclonal antibodies, or their portions that can bind to the protein of the invention. They also include antibodies of all classes. Furthermore, special antibodies such as humanized antibodies are also included.

[0128] The polyclonal antibody of the invention can be obtained according to the standard method by synthesizing an oligopeptide corresponding to the amino acid sequence and immunizing rabbits with the peptide (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wily & Sons, Section 11.12-11.13). The monoclonal antibody of the invention can be obtained according to the standard method by purifying the protein expressed in E. coli, immunizing mice with the protein, and producing a hybridoma cell by fusing the spleen cells and myeloma cells (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wily & Sons, Section 11.4-11.11).

[0129] The antibody binding to the protein of the present invention can be used for purification of the protein of the invention, and also for detection and/or diagnosis of the abnormalities of the expression and structure of the protein. Specifically, proteins can be extracted, for example, from tissues, blood, or cells, and the protein of the invention is detected by Western blotting, immunoprecipitation, or ELISA, etc. for the above purpose.

[0130] Furthermore, the antibody binding to the protein of the present invention can be utilized for treating the diseases that associates with the protein of the invention. If the antibodies are used for treating patients, human antibodies or humanized antibodies are preferable in terms of their low antigenicity. The human antibodies can be prepared by immunizing a mouse whose immune system is replaced with that of human ("Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice" Mendez M.J. et al. (1997) Nat. Genet. 15: 146-156). The humanized antibodies can be prepared by recombination of the hypervariable region of a monoclonal antibody (Methods in Enzymology (1991) 203: 99-121).

[0131] The present invention further relates to databases comprising at least a sequence of polynucleotide and/or protein, or a medium recorded in such databases, selected from the sequence data of the nucleotide and/or the amino acids indicated in Table 370. The term "database" means a set of accumulated information as machine-searchable and readable information of nucleotide sequence. The databases of the present invention comprise at least one of the novel nucleotide sequences of polynucleotide provided by the present invention. The databases of the present invention can consist of only the sequence data of the polynucleotide provided by the present invention or can comprise other information on nucleotide sequences of known full-length cDNAs or ESTs. The databases of the present invention can be comprised of not only the information on the nucleotide sequences but also the information on the gene functions revealed by the present invention. Additional information such as names of DNA clones carrying the full-length cDNAs can be recorded or linked together with the sequence data in the databases.

[0132] The database of the present invention is useful for gaining complete gene sequence information from partial sequence information of a gene of interest. The database of the present invention comprises nucleotide sequence information of full-length cDNAs. Consequently, by comparing the information in this database with the nucleotide sequence of a partial gene fragment yielded by differential display method or subtraction method, the information on the full-length nucleotide sequence of interest can be gained from the sequence of the partial fragment as a starting clue.

[0133] The sequence information of the full-length cDNAs constituting the database of the present invention contains not only the information on the complete sequences but also extra information on expression frequency of the genes

as well as homology of the genes to known genes and known proteins. Thus the extra information facilitates rapid functional analyses of partial gene fragments. Further, the information on human genes is accumulated in the database of the present invention, and therefore, the database is useful for isolating a human homologue of a gene originating from other species. The human homologue can be isolated based on the nucleotide sequence of the gene from the original species.

[0134] At present, information on a wide variety of gene fragments can be obtained by differential display method and subtraction method. In general, these gene fragments are utilized as tools for isolating the full-length sequences thereof. When the gene fragment corresponds to an already-known gene, the full-length sequence is easily obtained by comparing the partial sequence with the information in known databases. However, when there exists no information corresponding to the partial sequence of interest in the known databases, cDNA cloning should be carried out for the full-length cDNA. It is often difficult to obtain the full-length nucleotide sequence using the partial sequence information as an initial clue. If the full-length of the gene is not available, the amino acid sequence of the protein encoded by the gene remains unidentified. Thus the database of the present invention can contribute to the identification of full-length cDNAs corresponding to gene fragments, which cannot be revealed by using databases of known genes.

[0135] The invention is illustrated more specifically with reference to the following examples, but is not to be construed as being limited thereto.

EXAMPLE 1

Construction of a cDNA library by the oligo-capping method.

[0136] The NT-2 neuron progenitor cells (Stratagene), a teratocarcinoma cell line from human embryo testis, which can differentiate into neurons by treatment with retinoic acid were used. The NT-2 cells were cultured according to the rnanufacturer's instructions as follows.

(1) NT-2 cells were cultured without induction by retinoic acid treatment ((NT2RM1, NT2RM2, NT2RM4)).
(2) After cultured, NT-2 cells were induced by adding retinoic acid, and then were cultured for 48 hours (NT2RP1).
(3) After cultured, NT-2 cells were induced by adding retinoic acid, and then were cultured for 2 weeks (NT2RP2, NT2RP3, NT2RP4).

[0137] Also, the human brain neuroglioma cell line H4 (ATCC HTG-148) (BNGH41), human neuroblastoma cell line SK-N-MC (ATCC HTB-10) (SKNMC1), and human retinoblastoma cell line Y79 (ATCC HTB-18) (Y79AA1) were cultured according to the culture conditions described in the ATCC catalogue. The cells were harvested separately, and mRNA was extracted from each cell by the method described in the literature (Molecular Cloning 2nd edition. Sambrook J., Fritsch, E.F., and Maniatis T. (1989) Cold Spring Harbor Laboratory Press). Furthermore, poly(A)+RNA was purified from the mRNA using oligo-dT cellulose.

[0138] Similarly, human placenta (PLACE1, PLACE2, PLACE3), human ovary cancer tissue (OVARC1), tissues from human embryo at 10 weeks, which is enriched with head (HEMBA1), or body (HEMBB1), human mammary gland (MAMMA1), human thyroid gland (THYRO1) were used to extract mRNA by the method described in the literature (Molecular Cloning 2nd edition. Sambrook J., Fritsch, E.F., and Maniatis T. (1989) Cold Spring Harbor Laboratory Press). Furthermore, poly(A)+RNA was purified from the mRNA using oligo-dT cellulose.

[0139] Each poly(A)+RNA was used to construct a cDNA library by the oligo-capping method (Maruyama M. and Sugano S. (1994) Gene 138: 171-174). Using the Oligo-cap linker (SEQ ID NO: 2541) and the Oligo-dT primer (SEQ ID NO: 2542), bacterial alkaline phosphatase (BAP) treatment, tobacco acid phosphatase (TAP) treatment, RNA ligation, the first strand cDNA synthesis, and removal of RNA were performed as described in the reference (Suzuki and Kanno (1996) Protein Nucleic acid and Enzyme. 41: 197-201; Suzuki Y. et al. (1997) Gene 200: 149-156). Next, 5'- and 3'-PCR primers (SEQ ID NO: 2543, and 2544, respectively) were used for performing PCR to convert the cDNA into double stranded cDNA, which was then digested with SfiI. Then, the DraIII-cleaved pUC19FL3 vector (Figure 1; for NT2RM1, and NT2RP1), or the DraIII-cleaved pME18SFL3 (Figure 1) (GenBank AB009864, expression vector; for NT2RM2, NT2RM4, NT2RP2, NT2RP3, NT2RP4, BNGH41, SKNMC1, Y79AA1, PLACE1, PLACE2, PLACE3, OVARC1, HEMBA1, HEMBB1, MAMMA1, and THYRO1) was used for cloning the cDNA in an unidirectional manner, and cDNA libraries were obtained. Then, the nucleotide sequence of the 5'- and 3'- ends of the cDNA clones was analyzed with a DNA sequencer (ABI PRISM 377, PE Biosystems) after sequencing reactions were performed with the DNA sequencing reagents (Dye Terminator Cycle Sequencing FS Ready Reaction Kit, dRhodamine Terminator Cycle Sequencing FS Ready Reaction Kit, or BigDye Terminator Cycle Sequencing FS Ready Reaction Kit, from by PE Biosystems) according to the instructions. The data were compiled into a database.

[0140] The full-length-enriched cDNA libraries except those for NT2RM1 and NT2RP1 were constructed using eukaryotic expression vector pME18SFL3. The vector contains SRα promoter and SV40 small t intron in the upstream

of the cloning site, and SV40 polyA added signal sequence site in the downstream. As the cloning site of pME18SFL3 has asymmetrical DraIII sites, and the ends of cDNA fragments contain SfiI sites complementary to the DraIII sites, the cloned cDNA fragments can be inserted into the downstream of the SRα promoter unidirectionally. Therefore, clones containing full-length cDNA can be expressed transiently by introducing the obtained plasmid directly into COS cells. Thus, the clones can be analyzed very easily in terms of the proteins that are the gene products of the clones, or in terms of the biological

[0141] Herein, the cDNA libraries and the name of each clone are related as shown in Table 2. Therein, "xxxxxx" represents the clone number of six digits. Thus, the sequences are named by the library name, the clone number plus F- for the 5'-end, or R- for the 3'-end.

Table 2

library: clone	5'-end sequence	3'-end sequence
NT2RM1:		
NT2RM1xxxxxx	F-NT2RM1xxxxxx	
NT2RP1:		
NT2RP1xxxxxx	F-NT2RP1xxxxxx	
NT2RM2:		
NT2RM2xxxxxx	F-NT2RM2xxxxxx	R-NT2RM2xxxxxx
NT2RM4:		
NT2RM4xxxxxx	F-NT2RM4xxxxxx	R-NT2RM4xxxxxx
NT2RP2:		
NT2RP2xxxxxx	F-NT2RP2xxxxxx	R-NT2RP2xxxxxx
NT2RP3:		
NT2RP3xxxxxx	F-NT2RP3xxxxxx	R-NT2RP3xxxxxx
NT2RP4:		
NT2RP4xxxxxx	F-NT2RP4xxxxxx	R-NT2RP4xxxxxx
BNGH41:		
BNGH41xxxxxx	F-BNGH41xxxxxx	R-BNGH41xxxxxx
SKNMC1:		
SKNMC1xxxxxx	F-SKNMC1xxxxxx	R-SKNMC1xxxxxx
Y79AA1:		
Y79AA1xxxxxx	F-Y79AA1xxxxxx	R-Y79AA1xxxxxx
PLACE1:		
PLACE1xxxxxx	F-PLACE1xxxxxx	R-PLACE1xxxxxx
PLACE2:		
PLACE2xxxxxx	F-PLACE2xxxxxx	R-PLACE2xxxxxx
PLACE3:		
PLACE3xxxxxx	F-PLACE3xxxxxx	R-PLACE3xxxxxx
OVARC1:		
OVARC1xxxxxx	F-OVARC1xxxxxx	R-OVARC1xxxxxx
HEMBA1:		
HEMBA1xxxxxx	F-HEMBA1xxxxxx	R-HEMBA1xxxxxx
HEMBB1:		
HEMBB1xxxxxx	F-HEMBB1xxxxxx	R-HEMBB1xxxxxx
MAMMA1:		
MAMMA1xxxxxx	F-MAMMA1xxxxxx	R-MAMMA1xxxxxx
THYRO1:		

THYRO1xxxxxx F-THYRO1xxxxxx R-THYRO1xxxxxx

EXAMPLE 2

Estimation of the fullness ratio of the 5'-ends of the clones contained in the cDNA libraries constructed by the oligo-capping method.

[0142] The fullness ratio at the 5'-end sequences of the 59,823 clones in the human cDNA libraries constructed by the oligo-capping method was determined as follows. Of all the clones whose 5'-end sequences were found in those of known human mRNA in the public database, a clone was judged to be "full-length", if it had a longer 5'-end sequence than that of the known human mRNA, or, even though the 5'-end sequence was shorter, if it contained the translation initiation codon. A clone which did not contain the translation initiation codon was judged to be "non-full-length". The fullness ratio ((the number of full-length clones)/(the number of full-length and non-full-length clones)) at the 5'-end of the cDNA clones from each library was determined by comparing with the known human mRNA. As a result, the fullness ratio of the 5'-ends was 63.5%. It suggests that the human cDNA clones obtained by the described method have complete 5'-ends with high probability.

EXAMPLE 3

Assessment of the fullness ratio of the 5'-end of the cDNA by the ATGpr and the ESTiMateFL.

[0143] The ATGpr, developed by Salamov A.A., Nishikawa T., and Swindells M.B. in the Helix Research Institute, is a program for prediction of the translation start codon based on the characteristics of the sequences in the vicinity of the ATG codon. The results are shown with expectations that an ATG is a true start codon (0.05-0.94). When this program is applied to general cDNAs without considering whether or not the ATG codons in the cDNAs are the true initiation codons of the cDNAs, both the sensitivity and the specificity of the results are estimated at 66%. Here, the sensitivity means the ratio of the number of codons judged to be initiation codons by the program to the total number of true initiation codons, and the specificity means the ratio of the number of true initiation codons to the number of codons judged to be initiation codons by the program. In contrast, when the program was applied to the 5'-end sequences of the clones from the cDNA library that was obtained by the oligo-capping method and that had 65% fullness ratio, the sensitivity and specificity of evaluation of a full-length clone (clone containing the N-terminal end of ORF) were improved to 82-83% by selecting only clones having the ATGpr1 score 0.6 or higher.

[0144] Furthermore, the program was used to assess the fullness of 18,959 clones in the human cDNA libraries obtained here, which have 5'-ends matched to a known human mRNA. Briefly, the maximal ATGpr1 score of the clones was determined, and then their 5'-end sequence was compared with the known human mRNA to estimate whether the clone is full-length or not. The result was summarized in Table 3. Based on the knowledge that known mRNAs, in general, are highly expressed in the cell, the expression levels of genes having a low number in the EST hit, which represent mRNAs whose expression levels are relatively low were examined, and the result is shown in Table 4.

[0145] In the table, the number of full-length clones indicate that of clones containing the N-terminal end of ORF, and so does the number of non-full-length clones that of clones without the N-terminal end of ORF. The fullness ratio represents (the number of full-length clones)/(the number of full-length clones plus the number of non-full-length clones).

Table 3

The maximal ATGpr1 score and the fullness ratio of the 5'-end sequences of clones obtained from human cDNA libraries constructed by the oligo-capping method; clones having a matched 5'-end with that of a known human mRNA.

maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
>=0.70	11,193	9,346	83.5%
>=0.50	13,369	10,549	78.9%
>=0.30	15,489	11,340	73.2%

Table 3 (continued)

The maximal ATGpr1 score and the fullness ratio of the 5'-end sequences of clones obtained from human cDNA libraries constructed by the oligo-capping method; clones having a matched 5'-end with that of a known human mRNA.

maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
>=0.15	17,394	11,811	67.9%
>=0.00	18,959	12,046	63.5%

Table 4

The maximal ATGpr1 score and the fullness ratio of the 5'-end sequences of the clones obtained from human cDNA libraries constructed by the oligo-capping method; clones having 5 EST hits or less among the clones having a matched 5'-end with that of a known human mRNA.

maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
>=0.70	2,801	1,934	69.0%
>=0.50	3,683	2,393	65.0%
>=0.30	4,683	2,707	57.8%
>=0.15	5,559	2,890	52.0%
>=0.00	6,113	3,013	49.8%

[0146] The ESTiMateFL, developed by Nishikawa and Ota in the Helix Research Institute, is a method for the selection of a clone with high fullness ratio by comparing with the 5'-end or 3'-end sequences of ESTs in the public database.

[0147] By the method, a cDNA clone is judged presumably not to be full-length if there exist any ESTs which have longer 5'-end or 3'-end sequences than the clone. The method is systematized for high throughput analysis. A clone is judged to be full-length if the clone has a longer 5'-end sequence than ESTs in the public database. Even if a clone has a shorter 5'-end, the clone is judged to be full-length if the difference in length is within 50 bases, and otherwise judged not to be full-length, for convenience. In case of the 5'-end sequence of the clones which matches a known mRNA, about 80% of the sequences that were judged to be full-length by comparing with ESTs was judged to be full-length by estimating the 5'-end sequence, as well; about 80% of the sequences that were judged to be not full-length by comparing with ESTs was judged to be not full-length by estimating the 5'-end sequence, as well. The accuracy of the prediction by comparing cDNA clones with ESTs is improved with increasing number of ESTs to be compared. However, when only a limited number of ESTs are available, the reliability becomes low. Thus, the method is effective in excluding clones with high probability of being non-full-length, from the cDNA clones that is synthesized by the oligo-capping method and that have the 5'-end sequences with about 60 % fullness ratio. In particular, the ESTiMateFL is efficiently used to estimate the fullness ratio at the 3'-end sequence of cDNA of a human unknown mRNA which has a significant number of ESTs in the public database.

[0148] The 18,959 clones isolated from human cDNA libraries constructed by the oligo-capping method, which have the 5'-end sequence that matches a known human mRNA, were estimated by using the ATGpr and ESTiMateFL. Briefly, the 5'-end sequence that matches a known human mRNA of the respective clone was analyzed to obtain the maximal ATGpr1 score, and compared with the ORF of the known human mRNA that matches it to determine whether the clone is full-length or not. Then, the 5'-end sequence of the respective clone was analyzed by the ESTiMateFL to judge whether the clone is full-length or not. Specifically, the 5'-end sequences that match a known human mRNA of the 18,959 clones constructed by the oligo-capping method were compared with those of ESTs by the ESTiMateFL and the clones other than those that are not full-length were selected. Then, the selected clones were used to analyze the relationship between the ATGpr and the fullness ratio. The result was summarized in Table 5. Also, among the selected, the clones in which the number of the EST hit is not more than 5 were selected and analyzed. The result was summarized in Table 6, which represents the result of the analysis of mRNA with relatively low abundance.

[0149] In the Tables, the number of full-length clones, the number of non-full-length clones, and the fullness ratio indicate the number of the clones that contain the N-terminus of the ORF, the number of the clones that do not contain

the N-terminus of the ORF, and (the number of full-length clones)/(the number of full-length clones) plus (the number of non-full-length clones), respectively.

Table 5

The maximal ATGpr1 score and the fullness ratio of the 5'-end sequence in the clones isolated from human cDNA libraries constructed by the oligo-capping method, which have the 5'-end sequence that matches a known human mRNA, and also other than those being not full-length according to the comparison with ESTs.

maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
>=0.70	9,068	8,349	92.1%
>=0.50	10,345	9,318	90.1%
>=0.30	11,425	9,964	87.2%
>=0.15	12,254	10,335	84.3%
>=0.00	12,785	10,484	82.0%

Table 6

The maximal ATGpr1 score and the fullness ratio of the 5'-end sequence in the clones isolated from human cDNA libraries constructed by the oligo-capping method, which have the 5'-end sequence that matches a known human mRNA, and also other than those being not full-length according to the comparison with ESTs, in which the number of the EST hit is not more than 5.

maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
>=0.70	1,959	1,510	77.1%
>=0.50	2,469	1,821	73.8%
>=0.30	2,975	2,046	68.8%
>=0.15	3,368	2,164	64.3%
>=0.00	3,661	2,226	60.8%

[0150] According to the above results, it was found that, in case of using clones isolated from human cDNA libraries constructed by the oligo-capping method, the fullness ratio of the clones that have low score in the ATGpr can be improved by assessing their 5'-end sequence using the combination of the ATGpr and the ESTiMateFL. Therefore, the method was applied to select a cDNA clone with high fullness ratio.

EXAMPLE 4

Clustering of the 5'-end and 3'-end sequences of cDNA clones.

[0151] The 5'-end and 3'-end sequences of cDNA clones were obtained, and clustered separately. Briefly, data of the single pass sequencing of the determined 5'-end and 3'-end of cDNA clones was subjected to the BLAST search between the sequence data of all the clones synthesized in Example 1, and clones that are supposed to be originating from the same gene were clustered into a group. For the 5'-end sequence, those having the consensus sequence of 95% identity 300 base pairs or more are clustered into the same group. For the 3'-end sequence, those having the consensus sequence of 90% identity 200 base pairs or more are clustered into the same group. Among the clusters of the 5'-end and 3'-end sequences, the sequence having the longest lead was chosen as the representative sequence of the cluster (group).

EXAMPLE 5

Characterization of the representative sequences and the sequences of clones

[0152] Data of the 5'-end sequences of the representative sequences and clones was characterized by the following

methods:

(1) judging whether it is identical to the sequence of mRNA or ESTs from human by the BLAST search of the GenBank or SwissProt, and examining whether it is full-length by comparing with the sequences of known mRNA and ESTs from human.
(2) determining the ATGpr1 score using all the initiation codons contained within the 5'-end sequence by the ATGpr which predict fullness ratio.
(3) predicting the existence of the signal sequence using all the initiation codons contained within the 5'-end sequence by the PSORT which predict signal.
and,
(4) only with the 5'-end sequences of the representative sequences of the clusters, examining the keywords in the top hit data of the homology search of the SwissProt.

[0153] Data of the characterized representative sequences and clones was used for the final selection of the clones.

EXAMPLE 6

Identity to the human mRNA and human EST, and comparison of the 5'-end length.

[0154] The clones and the representative sequences of the clusters were judged to be identical to any human mRNA, if their 5'-end sequence has a region of 200 nucleotides or longer with 94% or more identity to the mRNA. The clones and the representative sequences of the clusters were judged to be identical to any human EST, if their 5'-end sequence has a region of 200 nucleotides or longer with 90% or more identity to the EST.

[0155] The clones and the representative sequences of the clusters were judged to be full-length in comparison with human mRNA, if their 5'-end sequence is longer than those of the mRNA, or it contains the translation initiation site. The clones and the representative sequences of the clusters were judged to be full-length in comparison with human EST in the database, if their 5'-end sequence is longer than those of the EST, or even though it is shorter, the difference in length between the two sequences is 50 nucleotides or less, for convenience. Otherwise, the clones and the representative sequences of the clusters were judged to be not full-length.

EXAMPLE 7

Prediction of the fullness ratio by the ATGpr.

[0156] The score in the ATGpr1 is the expectation to be full-length based on calculations, and the higher score reflects the higher fullness ratio as shown in Example 3. Further, the maximal ATGpr1 score represents the score obtained with all the initiation codons contained in the 5'-end sequence of the clones and the representative sequences, and are used for the characterization.

EXAMPLE 8

Prediction of the existence of a signal sequence by the PSORT.

[0157] Prediction of the existence of a signal sequence by the PSORT was performed on all of the amino acid sequences predicted from all the initiation codons in the 5'-end sequence of the clones and the representative sequences of the clusters. By analyzing the presence or absence of the sequence which is predicted to be a signal sequence, which is characteristics of the N-terminus of many secretory proteins, cDNA clones encoding a secretory protein or membrane protein were selected.

EXAMPLE 9

Prediction of the protein function by the BLAST search.

[0158] The 5'-end sequence of the representative sequences of the cluster was analyzed by the BLAST homology search of the SwissProt. The obtained top hit data was classified into those identical to the 5'-end representative sequence (identity was 90% or higher), those not identical to the 5'-end representative sequence (identity was 60% or lower, and compared sequence was not more than 25 nucleotides), and those similar to the 5'-end representative sequence (the rest of the data).

[0159] All the keywords in the SwissProt data corresponding to the top hit data were selected, and the 5'-end representative sequences were classified by the keywords relating with functions.
The keywords relating with a secretory protein or membrane protein are the followings:

growth factor,
cytokine,
hormone,
receptor,
G-protein coupled receptor,
ionic channel,
voltage-gated channel,
calcium channel,
extracellular matrix,
transmembrane, and
signal.

[0160] The keywords relating to glycoprotein is glycoprotein.
[0161] The keywords relating to signal transduction are the followings:

serine/threonine-protein kinase,
tyrosine-protein kinase, and
calmodulin-binding.

[0162] The keywords relating to transcription are the followings:

transcription regulation and activator,
transcription regulation and repressor, and
nuclear protein and repressor.

[0163] The keywords relating to diseases are disease mutation, and syndrome.
[0164] Many keywords overlapped in the respective group (receptor and transmembrane, for example), and some keywords overlapped in different groups (secretory or membrane, and diseases, etc.).

EXAMPLE 10

Selection of clones by characterization.

[0165] From the data obtained by the above characterization, clones encoding a novel secretory protein or membrane protein, or proteins with other predicted functions were selected by the combination of the ATGpr1 score and the prediction of the signal sequence by the PSORT, or according to the top hit data in the homology search of the SwissProt.
[0166] In selecting the clones, the 5'-end sequences that are identical to a human mRNA were ignored, whereas those that are identical to a human mRNA in part but obviously not identical in the other part were included. Because there were clones selected that are identical to a human mRNA in part but obviously not identical in the other part.
[0167] Also, if the finally selected clones were found to be not full-length compared with the sequences of human mRNA and ESTs, these clones were discarded.

EXAMPLE 11

A method for selection of clones by the combination of the ATGpr1 score and the prediction of the signal sequence by the PSORT (a method for selection of secretory proteins and membrane proteins that are novel and full-length).

[0168] The sequences of clones and the representative sequences of their clusters were used to obtain the maximal ATGpr1 score and predict the presence of the signal sequence. First, clones were selected based on the representative sequences of the clusters. The correspondence between the name and SEQ ID of the representative sequences used for selection (Table 368), and the correspondence between the name and SEQ ID of the introns (including the representative sequences of the 5'-end and 3'-end, and ESTs) used for selection of clones from the representative sequences of the groups (Table 369) were shown in the last part of the present specification. Therein, HRIFA and HRIRA indicate the representative sequence of the 5'-end group, and that of the 3'-end group, respectively.

[0169] In the clusters in which a single clone is contained (the sequence of the 5'-end clone = the representative sequence of the 5'-end), selected were the clones that were judged to be full-length in comparison with human mRNA and ESTs, having the maximal ATGpr1 score 0.5 or higher, and predicted to contain the signal sequence, in principle. However, in the following cases, a clone having a longer 5'-end was selected: the maximal ATGpr1 score was less than 0.5, the sequence of the 5'-end was not full-length, the clone was obviously shorter although the clone was not classified into the same cluster according to the BLAST search of the other clones, or the 5'-end sequence corresponding to the 3'-end of the other clones in the same cluster in which the 3'-end sequence of the clone was contained was found to be longer by assembling. Furthermore, if there were multiple full-length clones in the same cluster and it was not successful to determine by assembling which has the longer 5'-end, all the clones were selected. For assembling, the Sequencher(™) (Hitachi Soft Engineering) was used. As a result, the signal sequence predicted to be present in the representative sequence was not present in some of the selected clones. In some cases, the ATGpr1 score became smaller than 0.5 or 0.3. The fullness ratio in these clones was low, yet still it is possible that the clones are full-length. The clones in which the signal sequence predicted to be present in the representative sequence was not present after selection were likely to be without the signal sequence, but still it is possible that the clones encode a membrane protein.

[0170] In the clusters comprising multiple clones, in which the representative sequence of the 5'-end was predicted to contain the signal sequence, selected were the clones having the longest 5'-end sequence among the clones which were judged to be full-length compared with human mRNA and ESTs, having the maximal ATGpr1 score for the 5'-end sequence 0.5 or higher, and predicted to contain the signal sequence. However, in the following cases, a clone having a longer 5'-end was selected: the maximal ATGpr1 score was less than 0.5, the sequence of the 5'-end was not full-length, the clone was obviously shorter although the clone was not classified into the same cluster according to the BLAST search of the other clones, or the 5'-end sequence corresponding to the 3'-end of the other clones in the same cluster in which the 3'-end sequence of the clone was contained was found to be longer. Furthermore, if there were multiple full-length clones in the same cluster and it was not successful to determine by assembling which has the longer 5'-end, all the clones were selected. As a result, the signal sequence predicted to be present in the representative sequence was not present in some of the selected clones. In some cases, the ATGpr1 score became smaller than 0.5 or 0.3. The fullness ratio in these clones was low, yet still it is possible that the clones are full-length. The clones in which the signal sequence predicted to be present in the representative sequence was not present after selection were likely to be without the signal sequence at the 5'-end, but still it is possible that the clones encode a membrane protein.

[0171] Next, in the clusters comprising multiple clones, in which the representative sequence of the 5'-end was predicted to have no signal sequence, selected were the clones which were judged to be full-length compared with human mRNA and ESTs, having the maximal ATGpr1 score for the 5'-end sequence 0.5 or higher, and predicted to contain the signal sequence.

[0172] The number of the clones selected by the combination of the ATGpr1 score and the prediction of a signal sequence by the PSORT were 254. The number of the clones having the maximal ATGpr1 score 0.5 or higher, and predicted to contain a signal sequence were 170 (Table 7-10). Among the clones, 164 clones were found to have the representative sequence of the original cluster that fulfills the same conditions. On the other hand, 5 clones were selected from the representative sequences of the 5'-end of the clusters which was predicted to contain a signal sequence while the maximal ATGpr1 score was lower than 0.5. A clone was selected from the representative sequence of the 5'-end of the cluster which was predicted to have no signal sequence.

[0173] The clones that have the maximal ATGpr1 score 0.3 or higher and less than 0.5 and predicted to contain the signal sequence were 35 clones (Table 11), in which 8 clones were found to have the representative sequence of the original cluster that fulfills the same conditions. Twenty-seven clones were selected from the representative sequences of the clusters which have the maximal ATGpr1 score 0.5 or higher and were predicted to have no signal sequence.

[0174] The clones that have the maximal ATGpr1 score less than 0.3 and were predicted to contain a signal sequence were 41 clones (Table 12). The clones that have the maximal ATGpr1 score 0.5 or higher and were predicted to have no signal sequence were 4 clones (Table 13). The clones that have the maximal ATGpr1 score 0.3 or higher and less than 0.5 and were predicted to have no signal sequence were 2 clones (Table 14). The clones that have the maximal ATGpr1 score less than 0.3 and were predicted to contain a signal sequence were 2 clones (Table 15). The representative sequences of the original clusters of all the clones had the maximal ATGpr1 score 0.3 or higher, and were predicted to contain a signal sequence.

[0175] The fullness ratio of the clones having the maximal ATGpr1 score 0.5 or higher, 0.3 or higher, and 0 or higher is expected to be as shown in Table 3, 4, 5, and 6.

Table 7

Table 13

Four clones from which selected clones have the maximal ATGpr1 score 0.5 or higher, and predicted to be lacking the signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	name of representative sequence	maximal ATGpr1 score	signal
NT2RP3002281	F-NT2RP3002281	0.81	No	HRIFA012999a	0.61	Yes
NT2RP3002721	F-NT2RP3002721	0.94	No	HRIFA023305a	0.57	Yes
NT2RP3004083	F-NT2RP3004083	0.94	No	HRIFA008387a	0.76	Yes
PLACE1005669	F-PLACE1005669	0.94	No	HRIFA012513a	0.65	Yes

Table 14

Two clones from which selected clones have the maximal ATGpr1 score 0.3 or higher and less than 0.5 and predicted to have no signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	representative sequence	maximal ATGpr1 score	signal
NT2RP3000481	F-NT2RP3000481	0.47	No	HRIFA028614a	0.93	Yes
NT2RP3003559	F-NT2RP3003559	0.48	No	HRIFA025514a	0.45	Yes

Table 15

Two clones from which selected clones have the maximal ATGpr1 score 0 or higher and less than 0.3, and predicted to have no signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	representative sequence	maximal ATGpr1 score	signal
PLACE1005601	F-PLACE1005601	0.12	No	HRIFA010593a	0.64	Yes
PLACE1006786	F-PLACE1006786	0.22	No	HRIFA012333a	0.51	Yes

EXAMPLE 12

A method for the selection of clones based on the top hit data in the homology search against the SwissProt (a method for the selection of a novel full-length protein that is predicted to have a function based on the top hit data).

[0176] The representative sequences of the clusters were discarded in which the 5'-end sequence is identical (90% or more matching), or not similar (the compared part contains a sequence of 25 nucleotides or shorter and the similarity is lower than 60%) to the top hit data in the SwissProt. Then, the remaining representative sequences which has similarity to the representative sequences of the 5'-ends were classified by a group of the above keywords (some representative sequences belong to a group by multiple keywords), and then clones were selected from the clusters. The names and the corresponding SEQ IDs of the representative sequences, and also the names of the introns (including the representative sequence of the 5'-end or the 3'-end, or ESTs) used for selecting the clones from the representative sequences and the corresponding SEQ IDs are shown in the last part of the present specification (Table 368 and 369, respectively). HRIFA indicates the representative sequence of the 5'-end group, and HRIRA indicates the representative sequence of the 3'-end group.

[0177] In principle, from the clusters containing only a single clone (the 5'-end sequence is the representative sequence of the cluster), the clone was selected. However, in the following cases, the clone containing a longer 5'-end was selected: where the maximal ATGpr1 score was less than 0.5, the 5'-end sequence of the clone to be selected was not complete, or the 5'-end of the clone was found to be obviously short nevertheless the clone should not be

included in the same cluster based on the BLAST analysis between the other clones, or further, the 5'-end sequence of the said clone, which corresponds to the 3'-ends of the other clones belonging to the same cluster in which the 3'-end of the said clone was included, was turn out to be longer than those of the other clones by assembling them. When there were two clones in the same cluster, judged to be full-length, and it was difficult to determine which clone has the longer 5'-end even by assembling them, all the clones were selected. As a result, the ATGpr1 score in some clones became less than 0.5 or less than 0.3. The fullness ratio of these clones became lower, but there is still a possibility that the clones are full-length.

[0178] In the case in which multiple clones were contained in a cluster, selected was the clone having the longest 5'-end in the clones judged to be full-length compared to the human mRNA or human EST. However, in the following cases, the clone containing a longer 5'-end was selected: where the maximal ATGpr1 score was less than 0.5, the 5'-end sequence of the clone to be selected was not complete, or the 5'-end of the clone was found to be obviously short nevertheless the clone should not be included in the same cluster based on the BLAST analysis between the other clones, or further, the 5'-end sequence of the said clone, which corresponds to the 3'-ends of the other clones belonging to the same cluster in which the 3'-end of the said clone was included, was turn out to be longer than those of the other clones by assembling them. When there were two clones in the same cluster, judged to be full-length, and it was difficult to determine which clone has the longer 5'-end even by assembling them, all the clones were selected. As a result, the ATGpr1 score in some clones became less than 0.5 or less than 0.3. These clones can still be full-length.

[0179] Based on the top hit data in the SwissProt homology search, 658 clones were selected. Among them, 446 clones were selected by the keywords, secretion or membrane. Using the keyword, glycoprotein, 243 clones were selected. 51 clones were selected by the keywords for signal transduction. With the keywords for transcription, 130 clones were selected. 17 clones were selected by the keywords for disease.

[0180] Among the 446 clones selected by the keywords, secretion or membrane, 77 clones were overlapped with those selected by combining the ATGpr1 score and prediction by the PSORT for the existence of a signal sequence. Also, many clones were overlapped with those selected by the keyword, glycoprotein. Moreover, some clones were overlapped with the clones selected by the keywords for diseases.

[0181] Among the 243 clones selected by the keyword, glycoprotein, 53 clones were overlapped with those selected by combining the ATGpr1 score and prediction by the PSORT for the existence of a signal sequence. Also, many clones were overlapped with those selected by the keywords, secretion or membrane. Moreover, some clones were overlapped with the clones selected by the keywords in diseases.

[0182] Among the clones selected by the top hit data in the homology search on the SwissProt, 532 clones were having the maximal ATGpr1 score 0.5 or higher. 59 clones were having the maximal score 0.3 or higher and less than 0.5. 67 clones were with the maximal score less than 0.3.

[0183] When the maximal ATGpr1 score is 0.5 or higher, 0.3 or higher, no less than 0, the expected fullness ratio is as shown in Table 3, 4, 5, and 6, respectively.

Table 16

The representative sequences of the most homologous sequences in the SwissProt with the keyword(s) "growth factor", "cytokine", or "hormone", and the selected clones.

name of clone	name of representative sequence
HEMBA1001563	HRIFA001439a
HEMBA1003047	HRIFA002743a
HEMBA1005070	HRIFA020144a
HEMBA1006724	HRIFA021620a
HEMBA1006916	HRIFA021855a
MAMMA1001066	HRIFA027355a
MAMMA1001634	HRIFA027187a
MAMMA1002165	HRIFA027673a
NT2RM4000326	HRIFA032530a
NT2RM4001377	HRIFA005300a
NT2RP2000447	HRIFA006448a
NT2RP2000663	HRIFA006609a
NT2RP2000903	HRIFA006798a
NT2RP2002974	HRIFA027860a
NT2RP2003369	HRIFA008596a
NT2RP2004141	HRIFA009123a

Table 22

The representative sequences of the most homologous sequences in the SwissProt with the keyword(s) "serine/threonine-protein kinase", "tyrosine-protein kinase", or "calmodulin-binding", and the selected clones	
name of clone	name of representative sequence
HEMBA1001878	HRIFA001712a
HEMBA1002195	HRIFA017703a
HEMBA1002227	HRIFA019136a
HEMBA1002551	HRIFA002309a
HEMBA1005084	HRIFA020184a
HEMBA1005913	HRIFA029866a
HEMBA1005929	HRIFA020335a
HEMBB1000668	HRIFA029792a
MAMMA1000881	HRIFA026659a
MAMMA1001150	HRIFA026813a
MAMMA1002142	HRIFA027656a
NT2RM2000589	HRIFA021794a
NT2RM2001902	HRIFA031986a
NT2PP1001020	HRIFA005728a
NT2RP1001031	HRIFA005732a
NT2RP2001469	HRIFA028061a
NT2RP2001529	HRIFA007256a
NT2RP2001769	HRIFA007435a
NT2RP2003179	HRIFA008459a
NT2RP2003545	HRIFA008717a
NT2RP2004670	HRIFA028468a
NT2RP3000011	HRIFA022177a
NT2RP3000022	HRIFA022182a
NT2RP3000172	HRIFA022265a
NT2RP3000201	HRIFA022546a
NT2RP3000820	HRIFA018262a
NT2RP3003527	HRIFA025492a
NT2RP3003849	HRIFA025250a
NT2RP3003874	HRIFA025261a
NT2RP4000634	HRIFA029866a
NT2RP4000962	HRIFA027681a
OVARC1000255	HRIFA010975a
OVARC1000529	HRIFA011179a
OVARC1000916	HRIFA011449a
OVARC1001338	HRIFA019869a
PLACE1003135	HRIFA013726a
PLACE1005519	HRIFA015070a
PLACE1005736	HRIFA015453a
PLACE1008282	HRIFA016654a
PLACE1008297	HRIFA017031a
PLACE1010081	HRIFA018134a
PLACE1011364	HRIFA018904a
PLACE1011824	HRIFA019175a
THYRO1001205	HRIFA030237a
THYRO1001457	HRIFA030371a
THYRO1001593	HRIFA030448a

Table 25 (continued)

The clones selected by the keyword(s) of the top hit data in the SwissProt, and having the maximal score in the ATGpr1 0.5 or higher.		
name of clone	name of sequence	maximal ATGpr1 score
NT2RP2001662	F-NT2RP2001662	0.94
NT2RP2001755	F-NT2RP2001755	0.94
NT2RP2001769	F-NT2RP2001769	0.66
NT2RP2001878	F-NT2RP2001878	0.68
NT2RP2001921	F-NT2RP2001921	0.61
NT2RP2001948	F-NT2RP2001948	0.89
NT2RP2001956	F-NT2RP2001956	0.74
NT2RP2002063	F-NT2RP2002063	0.94
NT2RP2002188	F-NT2RP2002188	0.78
NT2RP2002232	F-NT2RP2002232	0.90
NT2RP2002304	F-NT2RP2002304	0.94
NT2RP2002409	F-NT2RP2002409	0.94
NT2RP2002527	F-NT2RP2002527	0.58
NT2RP2002533	F-NT2RP2002533	0.87
NT2RP2002564	F-NT2RP2002564	0.94
NT2RP2002942	F-NT2RP2002942	0.66
NT2RP2002976	F-NT2RP2002976	0.94
NT2RP2003042	F-NT2RP2003042	0.93
NT2RP2003179	F-NT2RP2003179	0.94
NT2RP2003210	F-NT2RP2003210	0.61
NT2RP2003302	F-NT2RP2003302	0.79
NT2RP2003369	F-NT2RP2003369	0.93
NT2RP2003390	F-NT2RP2003390	0.79
NT2RP2003469	F-NT2RP2003469	0.90
NT2RP2003545	F-NT2RP2003545	0.55
NT2RP2003593	F-NT2RP2003593	0.94
NT2RP2003655	F-NT2RP2003655	0.83
NT2RP2003664	F-NT2RP2003664	0.89
NT2RP2004069	F-NT2RP2004069	0.76
NT2RP2004108	F-NT2RP2004108	0.91
NT2RP2004141	F-NT2RP2004141	0.53
NT2RP2004447	F-NT2RP2004447	0.93
NT2RP2004606	F-NT2RP2004606	0.94
NT2RP2004648	F-NT2RP2004648	0.94
NT2RP2004670	F-NT2RP2004670	0.94
NT2RP2004794	F-NT2RP2004794	0.65
NT2RP2004847	F-NT2RP2004847	0.94
NT2RP2005069	F-NT2RP2005069	0.89
NT2RP2005163	F-NT2RP2005163	0.79
NT2RP2005181	F-NT2RP2005181	0.87
NT2RP2005247	F-NT2RP2005247	0.77
NT2RP2005425	F-NT2RP2005425	0.77
NT2RP2005535	F-NT2RP2005535	0.51
NT2RP2005597	F-NT2RP2005597	0.74
NT2RP2005632	F-NT2RP2005632	0.87
NT2RP2005666	F-NT2RP2005666	0.77
NT2RP2005774	F-NT2RP2005774	0.87

EXAMPLE 13

Selection of cDNA clone NT2RP2036580

[0184] Clone NT2RP2006580 as well as clone HEMBA1000121 was selected from the representative sequences belonging to HRIFA000116a cluster of the most homologous sequence in the SwissProt with the keywoods "transmembrane". Although each of the clones, HEMBA1000121 and NT2RP2006580, was assembled with other clones for 5' extension, any other clones did not extend the clones toward the 5' direction. Accordingly, it is possible that both clones are full-length cDNA clones. The maximal ATGpr1 score of F-NT2RP2006580 is 0.37, and therefore, the fullness ratio is low. However, it is still possible for the sequence to cover the full-length.

[0185] Thus, the total number of selected clones is 830. Based on the top matching data resulted from Swiss-Prot homology search, 659 clones were selected. From them, 447 clones were selected by the keywords of "secretion" and "membrane". Among the clones selected based on the top matching data, 60 clones exhibited the maximal ATGpr1 score of 0.3 or higher and less than 0.5.

[0186] The sequences of F-NT2RP2006580 and R-NT2RP2006580 are shown in SEQ ID NO: 2545 and SEQ ID NO: 2546, respectively.

EXAMPLE 14

Full-length sequence analysis and homology search

[0187] Full-length sequence was determined for each selected cDNA clones. The nucleotide sequence determination was performed mainly by the dye-terminator method using custom synthesized DNA primers according to the primer walking procedure (custom synthesized DNA primers were used for sequencing; sequencing reaction was performed with DNA sequencing reagent supplied by PE Biosystems according to the supplier's manual; and the samples were analyzed in an automatic sequencer made by the same supplier). Sequence determination of some clones was carried out in the same manner but using a Licor DNA sequencer. Overlapping partial nucleotide sequences, which were obtained by the above-described method, were assembled together to determine a full-length nucleotide sequence. Amino acid sequences were then deduced from the determined full-length nucleotide sequences. However, amino acid sequence is not shown for a clone of which coding region was hard to be deduced or of which amino acid sequence has less than 100 amino acid residues. SEQ ID NOs corresponding to the respective clones are indicated in Table 370.

[0188] GenBank, Swiss-Prot and UniGene were searched for the determined nucleotide sequences by BLAST analysis. Matching data of cDNA clone which exhibits higher homology and of which functions are easily predicted based on the nucleotide sequences and the deduced amino acid sequences are selected from the BLAST analysis matching data with P value of 10^{-4} or less. The matching data selected are listed herein. However, there are some clones that did not match the criteria for judgment and such matching data of BLAST analysis are not shown herein. The results of homology search indicated in the last part of this specification are as follows.

[0189] Homology search result 1: data obtained by the homology search of Swiss-Prot database for representative sequences of the 5'-end cluster

[0190] Homology search result 2: homology of representative sequences of the 5'-end cluster to the data in Swiss-Prot database; the P value is 10^{-10} or less

[0191] Homology search result 3: homology of representative sequences of the 5'-end cluster to the data in Swiss-Prot database; the P value is higher than 10^{-10} and 10^{-4} or less

[0192] Homology search result 4: homology of representative sequences of the 5'-end cluster to the data in Swiss-Prot database; the P value is higher than 10^{-4} and 1 or less

[0193] Homology search result 5: data obtained by the homology search of Swiss-Prot database for 5'-end sequences of cDNA clone

[0194] Homology search result 6: data obtained by the homology search of GenBank database (http:/lwww.ncbi.nlm.nih.gov/web/GenBank/) except for EST and STS sequence data for 5'-end sequences of cDNA clone

[0195] Homology search result 7: data obtained by the homology search of GenBank database (http://www.ncbi.nlm.nih.gov/web/GenBank/) except for EST and STS sequence data for 3'-end sequences of cDNA clone

[0196] Homology search result 8: data obtained by the homology search of Human UniGene database (http://www.ncbi.nlm.nih.gov/Unigene) for 5'-end sequences of cDNA clone

[0197] Homology search result 9: data obtained by the homology search of Human UniGene database (http://www.ncbi.nlm.nih.gov/Unigene) for 3'-end sequences of cDNA clone

[0198] Homology search result 10: result obtained by the homology search for full-length nucleotide sequences and deduced amino acid sequences

[0199] The P value indicates similarity between two sequences as a score by considering the probability that the two

sequences are accidentally similar. In general, as the value is lower, the similarity is higher. In general, as the value is lower, the homology is higher (Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.L. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410; Gish, W. & States, D.J. (1993) "Identification of protein coding regions by database similarity search." Nature Genet. 3:266-272).

Example 15. Gene expression analysis with hybridization using high density DNA filter

[0200] Nylon membrane for DNA spotting was prepared according to the following procedure. E. coli was cultured in each well of a 96-well plate (in a LB medium at 37. for 16 hours). A sample of each culture was suspended in 10 . 1 of sterile water in a well of a 96-well plate. The plate was heated at 100. for 10 minutes. Then, the boiled samples were analyzed by PCR. PCR was performed in a 20 .1 solution by using TaKaRa PCR Amplification Kit (Takara) according to the supplier's protocol. Primers used for the amplification of an insert cDNA in a plasmid were a pair of sequencing primers, ME761FW (5' tacggaagtgttacttctgc 3'/ SEQ ID NO: 3591) and ME1250RV (5' tgtgggaggttttttctcta 3'/ SEQ ID NO: 3592), or a pair of primers, M13M4 (5' gttttcccagtcacgac 3'/ SEQ ID NO: 3593) and M13RV (5' caggaaacagctatgac 3' / SEQ ID NO: 3594). PCR was performed using a thermal cycler, GeneAmp System 9600 (PE Biosystems) at 95. for 5 minutes; at 95. for 10 seconds and at 68. for 1 minute for 10 cycles; at 98. for 20 seconds and at 60. for 3 minutes for 20 cycles; and at 72. for 10 minutes. After the PCR, the 20 .1 reaction solution was loaded onto a 1% agarose gel and fractionated by electrophoresis. DNA on the gel was stained with ethidium bromide to confirm the amplification of cDNA. When cDNAs were not amplified by PCR, plasmids containing the corresponding insert cDNAs were prepared by the alkali-extraction method (J. Sambrook, E.F., Fritsh, & T. Maniatis, "Molecular Cloning, A laboratory manual/ 2nd edition, Cold Spring Harbor Laboratory Press, 1989).

[0201] Preparation of DNA array was carried out by the following procedure. A sample of a DNA solution was added in each well of a 384-well plate. DNA was spotted onto a nylon membrane (Boehringer) by using a 384-pin tool of Biomek 2000 Laboratory Automation System (Beckman-Coulter). Specifically, the 384-well plate containing the DNA was placed under the 384-pin tool. The independent 384 needles were simultaneously dipped into the DNA solution for DNA deposition. The needles were gently pressed onto a nylon membrane and the DNA deposited at the tips of needles was spotted onto the membrane. Denaturation of the spotted DNA and immobilization of the DNA on the nylon membrane were carried out according to standard methods (J. Sambrook, E.F., Fritsh, & T. Maniatis, "Molecular Cloning, A laboratory manual/ 2nd edition, Cold Spring Harbor Laboratory Press, 1989).

[0202] A probe for hybridization was radioisotope-labeled first strand cDNA. Synthesis of the first strand cDNA was performed by using Thermoscript(TM) RT-PCR System (GIBCO). Specifically, the first strand cDNA was synthesized by using 1.5 .g of mRNAs from various human tissues (Clontech), 1 .1 of 50.M Oligo(dT)20 and 50.Ci [.^{33}P]dATP according to an attached protocol. Purification of a probe was carried out by using ProbeQuant(TM) G-50 micro column (Amersham-Pharmacia Biotech) according to an attached protocol. In the next step, 2 units of E. coli RNase H were added to the reaction mixture. The mixture was incubated at room temperature for 10 minutes, and then, 100.g of human COT-1 DNA (GIBCO) was added thereto. The mixture was incubated at 97. for 10 minutes and then was allowed to stand on ice to give hybridization probe.

[0203] Hybridization of the radioisotope-labeled probe to the DNA array was performed according to standard methods (J. Sambrook, E.F., Fritsh, & T. Maniatis, Molecular Cloning, A laboratory manual/ 2nd edition, Cold Spring Harbor Laboratory Press, 1989). The membrane was washed as follows: the nylon membrane was washed 3 times by incubating it in Washing solution 1 (2xSSC, 1% SDS) at room temperature (about 26.) for 20 minutes; then the membrane was washed 3 times by incubating it in Washing solution 2 (0.1xSSC, 1% SDS) at 65. for 20 minutes.

[0204] Autoradiography was performed by using an image plate for BAS2000 (Fuji Photo Film Co., Ltd.). Specifically, the nylon membrane with probe hybridized thereon was wrapped with a piece of Saran Wrap and brought into tight contact with the image plate on the light-sensitive surface. The membrane with the image plate was placed in an imaging cassette for radioisotope and allowed to stand in dark place for 4 hours. The radioactivity recorded on the image plate was analyzed by using BAS2000 (Fuji Photo Film Co., Ltd.). The activity was subjected to electronic conversion and recorded as an image file of autoradiogram. The signal intensity of each DNA spot was analyzed by using Visage High Density Grid Analysis Systems (Genomic Solutions Inc.). The signal intensity was converted into numerical data. The data were taken in duplicate. The reproducibility was assessed by comparing the signal intensities of the corresponding spots on the duplicated DNA filters that were hybridized to a single DNA probe (Figure 2). In 95% of entire spots, the ratio between the corresponding spots falls within a range of 2 or less, and the correlation coefficient is r=1.97. Thus, the reproducibility is satisfactory.

[0205] The detection sensitivity in gene expression analysis was estimated by examining increases in the signal intensity of probe concentration-dependent spot in hybridization using a probe complementary to the DNA spotted on the nylon membrane. DNA used was PLACE 1008092 (the same as DNA deposited in GenBank under an Accession No. AF107253). The DNA array with DNA of PLACE1008092 was prepared according to the above-mentioned method. The probe used was prepared as follows: mRNA was synthesized in vitro from the clone, PLACE1008092. By using

this mRNA as a template, radioisotope-labeled first strand cDNA was synthesized in the same manner as described above, and the cDNA was used as the probe. In order to synthesize mRNA from PLACE1008092 in vitro, a plasmid in which the 5' end of the cDNA PLACE1008092 was ligated to the T7 promoter of pBluescript SK(-) was constructed. Specifically, the PLACE1008092 insert was cut out from pME18SFL3 carrying the cDNA at a DraIII site thereof by XhoI digestion. The resulting PLACE1008092 fragment was ligated to XhoI-predigested pBluescript SK(-) by using DNA ligation kit ver.2 (Takara). The in vitro mRNA synthesis from PLACE1008092 inserted into pBluescript SK(-) was carried out by using Ampliscribe(TM) T7 high yield transcription kit (Epicentre technologies). Hybridization and the analysis of signal intensity of each DNA spot were performed by the same methods as described above. When the probe concentration is $1x10^7$.g/ml or less, there was no increase of signal intensity proportional to the probe concentration. Therefore, it was assumed to be difficult to compare the signals with one another in this concentration range. Thus, the spots with the intensity of 40 or less were uniformly taken as low level signals (Figure 3). Within a concentration of the probe ranging from $1x10^7$.g/ml to 0.1.g/ml, the signal was found to increase in a probe concentration-dependent manner. The detection limit represented as the ratio of the expression level of test mRNA to that of total mRNA in a sample was 1:100,000.

[0206] Tables 28-184 (also containing clones without description in Examples) show the expression of each cDNA in human normal tissues (heart, lung, pituitary gland, thymus, brain, kidney, liver and spleen). The expression levels are indicated with numerical values of 0-10,000. Genes that were expressed in at least a single tissue are indicated below by the corresponding clone names:

clone: BNGH41000020, BNGH41000087, BNGH41000091, HEMBA1000121, HEMBA1000275, HEMBA1000300, HEMBA1000443, HEMBA1000462, HEMBA1000477, HEMBA1000634, HEMBA1000713, HEMBA1000835, HEMBA1000875, HEMBA1000940, HEMBA1000962, HEMBA1001228, HEMBA1001296, HEMBA1001390, HEMBA1001563, HEMBA1001621, HEMBA1002048, HEMBA1002131, HEMBA1002163, HEMBA1002164, HEMBA1002167, HEMBA1002178, HEMBA1002195, HEMBA1002227, HEMBA1002239, HEMBA1002316, HEMBA1002421, HEMBA1002524, HEMBA1002551, HEMBA1002767, HEMBA1002985, HEMBA1002992, HEMBA1003047, HEMBA1003072, HEMBA1003101, HEMBA1003120, HEMBA1003230, HEMBA1003294, HEMBA1003315, HEMBA1003392, HEMBA1003399, HEMBA1003487, HEMBA1003530, HEMBA1003945, HEMBA1004007, HEMBA1004067, HEMBA1001085, HEMBA1004110, HEMBA1004391, HEMBA1004444, HEMBA1004454, HEMBA1004505, HEMBA1004797, HEMBA1004952, HEMBA1005070, HEMBA1005084, HEMBA1005145, HEMBA1005230, HEMBA1005246, HEMBA1005337, HEMBA1005430, HEMBA1005449, HEMBA1005489, HEMBA1005545, HEMBA1005698, HEMBA1005929, HEMBA1005945, HEMBA1005016, HEMBA1006171, HEMBA1006276, HEMBA1006311, HEMBA1006335, HEMBA1006357, HEMBA1006430, HEMBA1006482, HEMBA1006517, HEMBA1006544, HEMBA1006658, HEMBA1006707, HEMBA1006749, HEMBA1006770, HEMBA1006902, HEMBA1006912, HEMBA1006916, HEMBA1006960, HEMBA1007013, HEMBA1007057, HEMBA1007063, HEMBA1007291, HEMBA1007332, HEMBB1000106, HEMBB1000309, HEMBB1000447, HEMBB1000542, HEMBB1000567, HEMBB1000642, HEMBB1000905, HEMBB1001026, HEMBB1001048, HEMBB1001407, HEMBB1001530, HEMBB1001573, HEMBB1001847, HEMBB1001959, HEMBB1001978, HEMBB1002039, HEMBB1002041, HEMBB1002051, HEMBB1002162, HEMBB1002228, HEMBB1002302, HEMBB1002427, HEMBB1002465, HEMBB1002661, HEMBB1002663, HEMBB1002693, MAMMA1000046, MAMMA1000102, MAMMA1000106, MAMMA1000118, MAMMA1000204, MAMMA1000226, MAMMA1000403, MAMMA1000449, MAMMA1000457, MAMMA1000473, MAMMA1000528, MAMMA1000591, MAMMA1000614, MAMMA1000652, MAMMA1000681, MAMMA1000706, MAMMA1000788, MAMMA1000810, MAMMA1000814, MAMMA1000881, MAMMA1000986, MAMMA1000994, MAMMA1001043, MAMMA1001066, MAMMA1001094, MAMMA1001141, MAMMA1001150, MAMMA1001284, MAMMA1001310, MAMMA1001344, MAMMA1001418, MAMMA1001532, MAMMA1001609, MAMMA1001615, MAMMA1001634, MAMMA1001893, MAMMA1001901, MAMMA1001957, MAMMA1002070, MAMMA1002091, MAMMA1002095, MAMMA1002128, MAMMA1002142, MAMMA1002165, MAMMA1002205, MAMMA1002224, MAMMA1002586, MAMMA1003126, NT2RM1000407, NT2RM1000462, NT2RM1000542, NT2RM1000789, NT2RM1000855, NT2RM1000858, NT2RM2000241, NT2RM2000306, NT2RM2000410, NT2RM2000423, NT2RM2000497, NT2RM2000514, NT2RM2000565, NT2RM2000582, NT2RM2000589, NT2RM2000622, NT2RM2000773, NT2RM2001126, NT2RM2001626, NT2RM2001792, NT2RM2001941, NT2RM4000198, NT2RM4000295, NT2RM4000444, NT2RM4000593, NT2RM4000761, NT2RM4000965, NT2RM4000997, NT2RM4001321, NT2RM4001325, NT2RM4001377, NT2RM4001735, NT2RM4001768, NT2RM4001843, NT2RP1000002, NT2RP1000181, NT2RP1000271, NT2RP1000300, NT2RP1000325, NT2RP1000465, NT2RP1000468, NT2RP1000740, NT2RP1000903, NT2RF1000981, NT2RP2000092, NT2RP2000178, NT2RP2000240, NT2RP2000447, NT2RP2000479, NT2RP2000533, NT2RP2000610, NT2RP2000616, NT2RP2000694, NT2RP2000739, NT2RP2001200, NT2RP2001223, NT2RP2001388, NT2RP2001469, NT2RP2001480, NT2RP2001514, NT2RP2001529, NT2RP2001538, NT2RP2001562, NT2RP2001662, NT2RP2001878, NT2RP2001903,

NT2RP2001921, NT2RP2001956, NT2RP2002015, NT2RP2002063, NT2RP2002188, NT2RP2002232, NT2RP2002409, NT2RP2002510, NT2RP2002527, NT2RP2002533, NT2RP2002564, NT2RP2002721, NT2RP2002824, NT2RP2002942, NT2RP2002974, NT2RP2003138, NT2RP2003179, NT2RP2003210, NT2RP2003302, NT2RP2003369, NT2RP2003383, NT2RP2003390, NT2RP2003469, NT2RP2003593, NT2RP2003599, NT2RP2003655, NT2RP2003940, NT2RP2003950, NT2RP2004069, NT2RP2004108, NT2RP2004141, NT2RP2004179, NT2RP2004205, NT2RP2004447, NT2RP2004524, NT2RP2004556, NT2RP2004606, NT2RP2004648, NT2RP2004794, NT2RP2004837, NT2RP2004847, NT2RP2005027, NT2RP2005069, NT2RP2005163, NT2RP2005181, NT2RP2005247, NT2RP2005378, NT2RP2005391, NT2RP2005425, NT2RP2005463, NT2RP2005535, NT2RP2005541, NT2RP2005597, NT2RP2005632, NT2RP2005666, NT2RP2005774, NT2RP2005878, NT2RP2005887, NT2RP2005941, NT2RP2006004, NT2RP2006042, NT2RP2006092, NT2RP2006099, NT2RP2006269, NT2RP3000011, NT2RP3000022, NT2RP3000059, NT2RP3000063, NT2RP3000125, NT2RP3000148, NT2RP3000171, NT2RP3000172, NT2RP3000201, NT2RP3000232, NT2RP3000304, NT2RP3000378, NT2RP3000436, NT2RP3000460, NT2RP3000645, NT2RP3000652, NT2RP3000676, NT2RP3000677, NT2RP3000721, NT2RP3000789, NT2RP3000818, NT2RP3000820, NT2RP3000838, NT2RP3000907, NT2RP3000921, NT2RP3001044, NT2RP3001159, NT2RP3001170, NT2RP3001195, NT2RP3001271, NT2RP3001388, NT2RP3001560, NT2RP3001592, NT2RP3001685, NT2RP3001738, NT2RP3001754, NT2RP3001858, NT2RP3001976, NT2RP3002015, NT2RP3002160, NT2RP3002281, NT2RP3002311, NT2RP3002324, NT2RP3002353, NT2RP3002409, NT2RP3002411, NT2RP3002721, NT2RP3002737, NT2RP3002738, NT2RP3002836, NT2RP3002900, NT2RP3002958, NT2RP3003000, NT2RP3003076, NT2RP3003354, NT2RP3003448, NT2RP3003469, NT2RP3003473, NT2RP3003532, NT2RP3003614, NT2RP3003729, NT2RP3003849, NT2RP3003874, NT2RP3003939, NT2RP3003963, NT2RP3004025, NT2RP3004067, NT2RP3004083, NT2RP3004090, NT2RP3004119, NT2RP3004130, NT2RP3004133, NT2RP3004202, NT2RP3004294, NT2RP3004309, NT2RP3004321, NT2RP3004355, NT2RP3004374, NT2RP3004406, NT2RP3004481, NT2RP3004552, NT2RP3004557, NT2RP3004625, NT2RP3004640, NT2RP3004647, NT2RP4000108, NT2RP4000634, NT2RP4001877, NT2RP4001879, NT2RP4002187, NT2RP4002715, NT2RP4002750, OVARC1000090, OVARC1000105, OVARC1000137, OVARC1000208, OVARC1000255, OVARC1000313, OVARC1000331, OVARC1000410, OVARC1000439, OVARC1000467, OVARC1000529, OVARC1000553, OVARC1000775, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000956, OVARC1000995, OVARC1001030, OVARC1001049, OVARC1001086, OVARC1001163, OVARC1001260, OVARC1001336, OVARC1001569, OVARC1001570, OVARC1001596, OVARC1001807, OVARC1001833, OVARC1001991, PLACE1000231, PLACE1000258, PLACE1000442, PLACE1000560, PLACE1000912, PLACE1000927, PLACE1001016, PLACE1001100, PLACE1001114, PLACE1001183, PLACE1001229, PLACE1001340, PLACE1001407, PLACE1001500, PLACE1001516, PLACE1001655, PLACE1001836, PLACE1001918, PLACE1002080, PLACE1002095, PLACE1002153, PLACE1002329, PLACE1002374, PLACE1002518, PLACE1002547, PLACE1002726, PLACE1002905, PLACE1002911, PLACE1002967, PLACE1003163, PLACE1003407, PLACE1003428, PLACE1003438, PLACE1003460, PLACE1003529, FLACE1003598, PLACE1003644, PLACE1003772, PLACE1003839, PLACE1003845, PLACE1003852, PLACE1004078, PLACE1004166, PLACE1004168, PLACE1004199, PLACE1004279, PLACE1004282, PLACE1004305, PLACE1004441, PLACE1004482, PLACE1004492, PLACE1004520, PLACE1004630, PLACE1004637, PLACE1004648, PLACE1004816, PLACE1004887, PLACE1005005, PLACE1005031, PLACE1005383, PLACE1005410, PLACE1005426, PLACE1005539, PLACE1005544, PLACE1005569, PLACE1005725, PLACE1005736, PLACE1005768, PLACE1005815, PLACE1005878, PLACE1005927, PLACE1006071, PLACE1006073, PLACE1006079, PLACE1006277, PLACE1006443, PLACE1006716, PLACE1006809, PLACE1007077, PLACE1007096, PLACE1007626, PLACE1007702, PLACE1008469, PLACE1008985, PLACE1009067, PLACE1009527, PLACE1009982, PLACE1010078, PLACE1010251, PLACE1010445, PLACE1011045, PLACE1011116, PLACE1011181, PLACE1011236, PLACE1011364, PLACE1011516, PLACE1011708, PLACE1011978, PLACE2000118, PLACE2000219, PLACE3000181, PLACE4000354, PLACE4000455, SKNMC1000014, THYRO1000061, THYRO1000099, THYRO1000584, THYRO1000795, THYRO1000866, THYRO1000999, THYRO1001063, THYRO1001113, THYRO1001128, THYRO1001205, THYRO1001237, THYRO1001242, THYRO1001456, THYRO1001457, THYRO1001478, THYRO1001495, THYRO1001523, THYRO1001529, THYRO1001593, THYRO1001608, THYRO1001700, THYRO1001702, THYRO1001725, THYRO1001770, THYRO1001803, Y79AA1000127, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000426, Y79AA1000521, Y79AA1000776, Y79AA1000777, Y79AA1000888, Y79AA1000967, Y79AA1001013, Y79AA1001090, Y79AA1001272, Y79AA1001328, Y79AA1001426, Y79AA1001427, Y79AA1001430, Y79AA1001523, Y79AA1001530, Y79AA1001592, Y79AA1001727, Y79AA1001787, Y79AA1001793, Y79AA1001799, Y79AA1001803, Y79AA1001863, Y79AA1002022, Y79AA1002213, Y79AA1002373, Y79AA1002376, Y79AA1002381.

PLACE1011116, PLACE1011181,
PLACE1011236, PLACE1011516, PLACE1011708, PLACE1011824, PLACE1011978, PLACE2000118,
PLACE3000181, SKNMC1000004, SKNMC1000014, THYRO1000584, THYRO1000866, THYRO1001113,
THYRO1001128, THYRO1001205, THYRO1001242, THYRO1001495, THYRO1001523, THYRO1001529,
THYRO1001593, THYRO1001608, THYRO1001702, THYRO1001725, THYRO1001770, THYRO1001803,
Y79AA1000117, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000426, Y79AA1000777,
Y79AA1000876, Y79AA1000888, Y79AA1000959, Y79AA1001013, Y79AA1001056, Y79AA1001090,
Y79AA1001264, Y79AA1001272, Y79AA1001328, Y79AA1001427, Y79AA1001430, Y79AA1001530,
Y79AA1001592, Y79AA1001727, Y79AA1001793, Y79AA1001799, Y79AA1001863, Y79AA1002022,
Y79AA1002213, Y79AA1002373, Y79AA1002376, Y79AA1002381.

[0216] Signal ratios of EC_AGE_BSA to EC_BSA and of EC_glycated_BSA to EC_BSA were calculated for each gene. Genes with high signal ratios were selected. In the case of calculating the ratio of signal value of 40 or less to that of more than 40, such signal values were, for convenience, taken as 40 instead of the real values. When the ratio EC_AGE_BSA/EC_BSA is 2 or more, expression of the genes exhibiting such ratio is expected to be elevated due to advanced glycation endproduct of bovine serum albumin. The higher the value is, the higher the gene expression level is. When the ratio EC_AGE_BSA/EC_BSA ranges from 0.5 to 2, expression of the genes exhibiting such ratio is expected to be unaffected due to advanced glycation endproduct of bovine serum albumin. When the ratio EC_AGE_BSA/EC_BSA is less than 0.5, expression of the genes exhibiting such ratio value is expected to be decreased due to advanced glycation endproduct of bovine serum albumin. The lower the value is, the lower the gene expression level is.

[0217] Clone with EC_AGE_BSA/BC_BSA ratio of 2 or higher are as follows: NT2RP2001538, NT2RP4001001 and Y79AA1000967.

[0218] These cDNAs are associated with diabetes.

Analysis of genes associated with neural cell differentiation

[0219] Genes involved in neural cell differentiation are useful for treating neurological diseases. It is possible that genes with varying expression levels in response to induction of cellular differentiation in neural cells are associated with neurological diseases.

[0220] A survey was performed for genes of which expression levels are varied in response to induction of differentiation (stimulation by retinoic acid (RA)) in cultured cells of a neural strain, NT2.

[0221] The NT2 cells were treated basically according to supplier's instruction manual. "Undifferentiated NT2 cells" means NT2 cells successively cultured in an Opti-MEM I (GIBCO-BRL; catalog No. 31985) containing 10%(v/v) fetal bovine serum and 1%(v/v) penicillin-streptomycin (GIBCO BRL). "NT2 cells cultured in the presence of retinoic acid" means the cells resulted from transferring undifferentiated NT2 cells into a retinoic acid-containing medium, which consists of D-MEM (GIBCO BRL; catalog No. 11965), 10%(v/v) fetal bovine serum, 1%(v/v) penicillin-streptomycin and 10.M retinoic acid (GIBCO-BRL), and the subsequent successive culture therein for 5 weeks. "NT2 cells that were cultured in the presence of retinoic acid and then further cultured in the presence of cell-division inhibitor added" means NT2 cells resulted from transferring NT2 cells cultured in the presence of retinoic acid for 5 weeks into a cell-division inhibitor-containing medium, which consisted of D-MEM(GIBCO BRL; catalog No.11965), 10%(v/v) fetal bovine serum, 1%(v/v) penicillin-streptomycin, 10. M retinoic acid, 10.M FudR (5-fluoro-2'-deoxyuridine: GIBCO BRL), 10. M Urd (Uridine: GIBCO BRL) and 1.M araC (Cytosine.-D-Arabinofuranoside: GIBCO BRL), and the subsequence successive culture for 2 weeks. Each of the cells were treated with trypsin and then harvested. Total RNAs were extracted from the cells by using S.N.A.P.(TM) Total RNA Isolation kit (Invitrogen). The labeling of probe used for hybridization was carried out by using 10.g of the total RNA according to the same methods as described above. The data were obtained in triplicate (n=3). The data of signal value representing gene expression level in the cells in the presence of stimulation for inducing differentiation were compared with those in the absence of the stimulation. The comparison was performed by statistical treatment of two-sample t-test. Clones with significant difference in the signal distribution were selected under the condition of $p<0.05$. In this analysis, clones with the difference can be statistically detected even when the signals were low. Accordingly, clones with signal value of 40 or less were also assessed for the selection.

[0222] Tables 186-365 show the expression level of each cDNA in undifferentiated NT2 cells, NT2 cells cultured in the presence of RA, and NT2 cells that were cultured in the presence of RA and that were further cultured in the presence of cell-division inhibitor added.

[0223] Averaged signal values (M_1, M_2) and sample variances (s_1^2, s_2^2) were calculated for each gene in each of the cells, and then, the pooled sample variances s^2 were obtained from the sample variances of the two types of cells to be compared. The t values were determined according to the following formula: $t=(M1-M2)/s/(1/3+1/3)^{1/2}$. When the determined t-value was greater than a t-value at P, which means the probability of significance level, of 0.05 or 0.01 in the t-distribution table with 4 degrees of freedom, the difference was judged to be found in the expression level of the gene between the two types of cells at $p<0.05$ or $p<0.01$, respectively. The tables also include the information on

an increase (+) or decrease (-) in the expression level of a gene in the treated cells when the level is compared with that of untreated undifferentiated cells.

[0224] Clones of which expression levels increased by RA are as follows: HEMBA1000121, HEMBA1000275, HEMBA1000300, HEMBA1000634, HEMBA1000671, HEMBA1000875, HEMBA1001184, HEMBA1001390, HEMBA1001886, HEMBA1002163, HEMBA1002227, HEMBA1002420, HEMBA1002421, HEMBA1003072, HEMBA1003120, HEMBA1003294, HEMBA1003497, HEMBA1004007, HEMBA1004110, HEMBA1004391, HEMBA1004444, HEMBA1005230, HEMBA1005246, HEMBA1005267, HEMBA1005489, HEMBA1005913, HEMBA1006299, HEMBA1006357, HEMBA1006517, HEMBA1006544, HEMBA1006658, HEMBA1006749, HEMBA1007063, HEMBA1007241, HEMBB1000447, HEMBB1000542, HEMBB1000567, HEMBB1000642, HEMBB1000668, HEMBB1001026, HEMBB1001847, HEMBB1002051, HEMBB1002120, HEMBB1002228, HEMBB1002693, MAMMA1000106, MAMMA1000141, MAMMA1000473, MAMMA1000528, MAMMA1000810, MAMMA1000881, MAMMA1001634, MAMMA1001957, MAMMA1002205, MAMMA1002224, NT2RM2000423, NT2RM2000497, NT2RM2000582, NT2RM2001126, NT2RM2001902, NT2RM4000198, NT2RM4000284, NT2RM4000593, NT2RM4001321, NT2RP1000002, NT2RP1000050, NT2RP1000181, NT2RP1000261, NT2RP1000465, NT2RP1000468, NT2RP1000579, NT2RP1000679, NT2RP2000092, NT2RP2000479, NT2RP2000610, NT2RP2000663, NT2RP2000694, NT2RP2000903, NT2RP2001388, NT2RP2001538, NT2RP2001878, NT2RP2001015, NT2RP2002304, NT2RP2002721, NT2RP2002824, NT2RP2002942, NT2RP2002974, NT2RP2002976, NT2RP2003179, NT2RP2003302, NT2RP2003383, NT2RP2003469, NT2RP2003664, NT2RP2003940, NT2RP2004069, NT2RP2004108, NT2RP2004524, NT2RP2004556, NT2RP2004670, NT2RP2005069, NT2RP2005247, NT2RP2005425, NT2RP2005463, NT2RP2005514, NT2RP2005535, NT2RP2005541, NT2RP2005774, NT2RP2005878, NT2RP2005883, NT2RP2005887, NT2RP2006099, NT2RP2006134, NT2RP3000011, NT2RP3000125, NT2RP3000171, NT2RP3000232, NT2RP3000460, NT2RP3000481, NT2RP3000652, NT2RP3000677, NT2RP3000818, NT2RP3000820, NT2RP3001044, NT2RP3001061, NT2RP3001170, NT2RP3001240, NT2RP3001322, NT2RP3001388, NT2RP3001542, NT2RP3001592, NT2RP3001976, NT2RP3002790, NT2RP3002900, NT2RP3002983, NT2RP3003000, NT2RP3003354, NT2RP3003532, NT2RP3003729, NT2RP3003874, NT2RP3003939, NT2RP3004025, NT2RP3004083, NT2RP3004090, NT2RP3004130, NT2RP3004202, NT2RP3004294, NT2RP3004640, NT2RP4000108, NT2RP4000634, NT2RP4002451, NT2RP4002715, OVARC1000090, OVARC1000208, OVARC1000275, OVARC1000553, OVARC1000775, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000995, OVARC1001030, OVARC1001049, OVARC1001132, OVARC1001596, OVARC1002178, PLACE1000258, PLACE1000442, PLACE1000927, FLACE1000986, PLACE1001100, PLACE1001123, PLACE1001795, PLACE1002518, PLACE1002547, PLACE1002967, PLACE1003407, PLACE1003428, PLACE1003644, PLACE1003839, PLACE1004078, PLACE1004441, PLACE1004450, PLACE1005669, PLACE1005682, PLACE1005736, PLACE1005768, PLACE1005815, PLACE1006073, PLACE1006208, PLACE1007296, PLACE1007626, PLACE1008282, PLACE1008984, PLACE1008985, PLACE1010445, PLACE1011708, PLACE1011978, PLACE4000455, SKNMC1000004, THYRO1000036, THYRO1000580, THYRO1000776, THYRO1000999, THYRO1001063, THYRO1001128, THYRO1001205, THYRO1001327, THYRO1001523, THYRO1001725, THYRO1001770, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1001056, Y79AA1001062, Y79AA1001090, Y79AA1001727, Y79AA1002213, Y79AA1002381.

[0225] Clones of which expression levels decreased by RA are as follows: BNGH41000020, HEMBA1005070, NT2RP2005027, NT2RP3003473, Y79AA1002376.

[0226] Clones of which expression levels increase by RA/inhibitor are as follows: HEMBA1000128, HEMBA1000875, HEMBA1001390, HEMBA1002163, HEMBA1002227, HEMBA1002421, HEMBA1004391, HEMBA1004454, HEMBA1004785, HEMBA1005913, HEMBA1006171, HEMBA1006299, HEMBA1006335, HEMBA1006544, HEMBA1007241, HEMBB1000447, HEMBB1000668, MAMMA1000994, MAMMA1001344, NT2RM2000582, NT2RP1001004, NT2RP2000663, NT2RP2000694, NT2RP2000903, NT2RP2001388, NT2RP2002674, NT2RP2002974, NT2RP2003383, NT2RP2004069, NT2RP2004606, NT2RP2004837, NT2RP2005069, NT2RP2005425, NT2RP2005463, NT2RP2005541, NT2RP2005883, NT2RP2005887, NT2RP3000460, NT2RP3000838, NT2RP3001044, NT2RP3001240, NT2RP3001388, NT2RP3002721, NT2RP3002738, NT2RP3003469, NT2RP3004083, NT2RP3004130, NT2RP3004202, NT2RP3004294, NT2RP3004640, NT2RP4000108, NT2RP4002451, NT2RP4002715, OVARC1000275, OVARC1000467, OVARC1000553, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000995, OVARC1001030, OVARC1001222, OVARC1001596, OVARC1002058, OVARC1002178, PLACE1000927, PLACE1001123, PLACE1001407, PLACE1001464, PLACE1001564, PLACE1001795, PLACE1002547, PLACE1003407, PLACE1003644, PLACE1003845, PLACE1004441, PLACE1004482, PLACE1005410, PLACE1005601, PLACE1005725, PLACE1005736, PLACE1006093, PLACE1006219, PLACE1006290, PLACE1006716, PLACE1007296, PLACE1007626, PLACE1008359, PLACE1010968, PLACE1011364,

HEMBB1001407, HEMBB1001573, HEMBB1001978, HEMBB1002041, HEMBB1002162, HEMBB1002245, HEMBB1002427, HEMBB1002693, MAMMA1000102, MAMMA1000106, MAMMA1000118, MAMMA1000141, MAMMA1000204, MAMMA1000226, MAMMA1000457, MAMMA1000473, MAMMA1000591, MAMMA1000681, MAMMA1000810, MAMMA1000986, MAMMA1001043, MAMMA1001141, MAMMA1001237, MAMMA1001344, MAMMA1001893, MAMMA1001957, MAMMA1001978, MAMMA1002070, MAMMA1002091, MAMMA1002095, MAMMA1002633, NT2RM1000580, NT2RM1000855, NT2RM1000858, NT2RM2000410, NT2RM2000565, NT2RM2001626, NT2RM2001939, NT2RM2001941, NT2RM4000444, NT2RM4000587, NT2RM4000648, NT2RM4000997, NT2RM4001325, NT2RM4001735, NT2RM4001768, NT2RM4002352, NT2RP1000050, NT2RP1000181, NT2RP1000261, NT2RP1000300, NT2RP1000448, NT2RP1000551, NT2RP1000613, NT2RP1000981, NT2RP1001563, NT2RP2000479, NT2RP2000533, NT2RP2000649, NT2RP2000663, NT2RP2000694, NT2RP2000818, NT2RP2000903, NT2RP2001200, NT2RP2001276, NT2RP2001495, NT2RP2001915, NT2RP2001956, NT2RP2002188, NT2RP2002232, NT2RP2002527, NT2RP2002533, NT2RP2002721, NT2RP2002824, NT2RP2002942, NT2RP2002976, NT2RP2003042, NT2RP2003390, NT2RP2003469, NT2RP2003593, NT2RP2003655, NT2RP2003664, NT2RP2003950, NT2RP2004179, NT2RP2004205, NT2RP2004495, NT2RP2004524, NT2RP2004556, NT2RP2004606, NT2RP2004648, NT2RP2004794, NT2RP2005163, NT2RP2005181, NT2RP2005463, NT2RP2005597, NT2RP2005666, NT2RP2005883, NT2RP2005994, NT2RP2006004, NT2RP2006269, NT2RP2006512, NT2RP2006580, NT2RP3000169, NT2RP3000171, NT2RP3000304, NT2RP3000460, NT2RP3000616, NT2RP3000721, NT2RP3000818, NT2RP3000907, NT2RP3000921, NT2RP3001159, NT2RP3001195, NT2RP3001240, NT2RP3001271, NT2RP3001322, NT2RP3001388, NT2RP3001560, NT2RP3001592, NT2RP3001650, NT2RP3001738, NT2RP3002015, NT2RP3002311, NT2RP3002342, NT2RP3002411, NT2RP3002790, NT2RP3002836, NT2RP3002900, NT2RP3002958, NT2RP3003000, NT2RP3003354, NT2FP3003532, NT2RP3003535, NT2RP3003614, NT2RP3004025, NT2RP3004075, NT2RP3004083, NT2RP3004090, NT2RP3004130, NT2RP3004294, NT2RP3004309, NT2RP3004345, NT2RP3004406, NT2RP3004481, NT2RP3004552, NT2RP4001001, NT2RP4001009, NT2RP4001467, NT2RP4001879, NT2RP4002187, NT2RP4002451, NT2RP4002750, OVARC1000003, OVARC1000105, OVARC1000307, OVARC1000439, OVARC1000553, OVARC1001030, OVARC1001336, OVARC1001570, PLACE1000231, PLACE1000560, PLACE1000740, PLACE1000912, PLACE1000914, PLACE1000927, PLACE1001016, PLACE1001183, PLACE1001231, PLACE1001401, PLACE1001407, PLACE1001464, PLACE1001536, PLACE1001564, PLACE1001655, PLACE1001836, PLACE1001918, PLACE1001949, PLACE1002518, PLACE1002726, PLACE1002967, PLACE1003573, PLACE1003737, PLACE1003839, PLACE1003845, PLACE1003852, PLACE1004279, PLACE1004282, PLACE1004441, PLACE1004637, PLACE1004648, PLACE1004816, PLACE1004887, PLACE1005003, PLACE1005005, PLACE1005410, PLACE1005544, PLACE1005569, PLACE1005660, PLACE1005725, PLACE1005745, PLACE1005927, PLACE1006290, PLACE1006443, PLACE1006959, PLACE1007096, PLACE1007296, PLACE1007626, PLACE1007881, PLACE1008359, PLACE1008469, PLACE1008716, PLACE1008985, PLACE1009196, PLACE1009279, PLACE1009527, PLACE1009546, PLACE1009600, PLACE1010011, PLACE1010078, PLACE1010445, PLACE1010713, PLACE1010784, PLACE1010968, PLACE1011236, PLACE1011516, PLACE3000181, THYRO1000400, THYRO1000678, THYRO1000776, THYRO1000956, THYRO1001102, THYRO1001113, THYRO1001205, THYRO1001237, THYRO1001242, THYRO1001266, THYRO1001327, THYRO1001478, THYRO1001523, THYRO1001641, THYRO1001702, THYRO1001725, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000521, Y79AA1000888, Y79AA1001013, Y79AA1001212, Y79AA1001264, Y79AA1001328, Y79AA1001426, Y79AA1001427, Y79AA1001727, Y79AA1001787, Y79AA1001795, Y79AA1001803, Y79AA1002058, Y79AA1002129, Y79AA1002213, Y79AA1002373,

[0242] Names of clones whose deduced amino acid sequences were predicted to have functional domains by Pfam search, and names of the matched functional domains are shown below. When multiple functional domains matched a clone, each domain name was indicated, separated by a double-slash mark,//.

HEMBA1000006//Src homology domain 3
HEMBA1000128//SCP-like extracellular Proteins
HEMBA1000349//ABC transporters
HEMBA1000462//RNA recognition motif. (aka RRM, RBD, or RNP domain)
HEMBA1000590//EGF-like domain//von Willebrand factor type A domain
HEMBA1000671//Zinc finger, C2H2 type
HEMBA1000732//EGF-like domain
HEMBA1000940//Connexin
HEMBA1001221//EGF-like domain//Kazal-type serine protease inhibitor domain

NT2RM4000444//ABC transporters
NT2RM4001377//PH (pleckstrin homology) domain
NT2RM4001768//Alcohol/other dehydrogenases, short chain type
NT2RM4002352//Low-density lipoprotein receptor domain class A
NT2RP1000181//Heme-binding domain in cytochrome b5 and oxidoreductases
NT2RP1000271//Zinc finger, C2H2 type
NT2RP1000325//Mitochondrial carrier proteins
NT2RP1000613//Eukaryotic-type carbonic anhydrases
NT2RP1000981//IG superfamily
NT2RP1001004//Thrombospondin type 1 domain
NT2RP1001020//Eukaryotic protein kinase domain
NT2RP1001031//WD domain, G-beta repeats
NT2RP1001563//EGF-like domain//Lectin C-type domain short and long forms//SCP-like extracellular Proteins
NT2RP2000092//Zinc finger, C2H2 type
NT2RP2000514//Fibronectin type III domain//IG superfamily
NT2RP2000649//Zinc-binding metalloprotease domain
NT2RP2000712//Zinc finger, C2H2 type
NT2RP2000739//Zinc finger, C2H2 type
NT2RP2001514//E1-E2 ATPases
NT2RP2001529//Eukaryotic protein kinase domain
NT2RP2001755//Thrombospondin type 1 domain
NT2RP2001769//Eukaryotic protein kinase domain
NT2RP2002188//Carboxylesterases
NT2RP2002527//Heme-binding domain in cytochrome b5 and oxidoreductases
NT2RP2002564//Zinc finger, C2H2 type
NT2RP2002942//IG superfamily
NT2RP2003179//Eukaryotic protein kinase domain
NT2RP2003302//Zinc finger, C2H2 type
NT2RP2003390//DnaJ, prokaryotic heat shock protein
NT2RP2003469//Sugar (and other) transporters
NT2RP2003545//Eukaryotic protein kinase domain
NT2RP2003593//Thioredoxins
NT2RP2003940//Zinc finger, C2H2 type
NT2RP2004108//Zinc finger, C2H2 type
NT2RP2004205//IG superfamily
NT2RP2004670//Eukaryotic protein kinase domain
NT2RP2004847//Zinc finger, C2H2 type
NT2RP2005181//Amino acid permeases
NT2RP2005247//Zinc finger, C3HC4 type (RING finger)
NT2RP2005391//Fibronectin type III domain
NT2RP2005535//Zinc finger, C2H2 type
NT2RP2005774//Zinc finger, C2H2 type
NT2RP2005878//Alcohol/other dehydrogenases, short chain type
NT2RP2005941//Homeobox domain//'Paired box' domain
NT2RP2006004//Fibronectin type III domain
NT2RP3000011//WD domain, G-beta repeats
NT2RP3000022//Eukaryotic protein kinase domain
NT2RP3000063//Zinc finger, C2H2 type
NT2RP3000148//Zinc finger, C2H2 type
NT2RP3000172//Eukaryotic protein kinase domain
NT2RP3000201//Eukaryotic protein kinase domain
NT2RP3000232//Zinc finger, C2H2 type
NT2RP3000304//Low-density lipoprotein receptor domain class A//Low-density lipoprotein receptor domain class B
NT2RP3000436//Thioredoxins
NT2RP3000460//eubacterial secY protein
NT2RP3000616//Fibronectin type III domain
NT2RP3000652//Zinc finger, C2H2 type

THYRO1000099, THYRO1000196, THYRO1000400, THYRO1000584, THYRO1000678, THYRO1000776, THYRO1000795, THYRO1000956, THYRO1001102, THYRO1001113, THYRO1001205, THYRO1001237, THYRO1001242, THYRO1001266, THYRO1001327, THYRO1001456, THYRO1001478, THYRO1001523, THYRO1001529, THYRO1001641, THYRO1001702, THYRO1001725, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000426, Y79AA1000521, Y79AA1000876, Y79AA1000888, Y79AA1000959, Y79AA1001013, Y79AA1001212, Y79AA1001264, Y79AA1001328, Y79AA1001426, Y79AA1001427, Y79AA1001430, Y79AA1001727, Y79AA1001787, Y79AA1001795, Y79AA1001799, Y79AA1001803, Y79AA1002022, Y79AA1002058, Y79AA1002129, Y79AA1002213, Y79AA1002373,

[0251] The following 146 clones were categorized into glycoprotein-associated proteins.

BNGH41000087, BNGH41000091, HEMBA1000349, HEMBA1000590, HEMBA1000745, HEMBA1000835, HEMBA1001221, HEMBA1001228, HEMBA1001621, HEMBA1002131, HEMBA1002178, HEMBA1002421, HEMBA1002767, HEMBA1003230, HEMBA1003392, HEMBA1004250, HEMBA1004391, HEMBA1004444, HEMBA1004505, HEMBA1005449, HEMBA1005522, HEMBA1005545, HEMBA1006707, HEMBA1006749, HEMBA1006902, HEMBB1000679, HEMBB1000881, HEMBB1001048, HEMBB1002120, HEMBB1002245, HEMBB1002427, MAMMA1000102, MAMMA1000591, MAMMA1000681, MAMMA1001043, MAMMA1001237, MAMMA1002070, MAMMA1002586, MAMMA1003126, NT2RM1000462, NT2RM1000580, NT2RM2001792, NT2RM2001818, NT2RM2001939, NT2RM2001941, NT2RM4000198, NT2RM4000284, NT2RM4000417, NT2RM4000648, NT2RM4000997, NT2RM4001325, NT2RM4002352, NT2RP1000613, NT2RP1000981, NT2RP1001004, NT2RP2000616, NT2RP2000694, NT2RP2000903, NT2RP2001480, NT2RP2001755, NT2RP2002533, NT2RP2003042, NT2RP2003210, NT2RP2004205, NT2RP2004606, NT2RP2005027, NT2RP2005181, NT2RP2005541, NT2RP2005597, NT2RP2005883, NT2RP2006004, NT2RP2006042, NT2RP2006269, NT2RP3000304, NT2RP3000616, NT2RP3000921, NT2RP3001650, NT2RP3002160, NT2RP3002737, NT2RP3002958, NT2RP3003000, NT2RP3003532, NT2RP3004130, NT2RP3004133, NT2RP3004481, NT2RP3004552, NT2RP3004640, NT2RP4000108, NT2RP4001467, NT2RP4002750, OVARC1000003, OVARC1000553, OVARC1000811, OVARC1000873, OVARC1001336, OVARC1001607, OVARC1001991, PLACE1000033, PLACE1000740, PLACE1001016, PLACE1001123, PLACE1001231, PLACE1001464, PLACE1001655, PLACE1001836, PLACE1002355, PLACE1002374, PLACE1002905, PLACE1002911, PLACE1003573, PLACE1003737, PLACE1003772, PLACE1003839, PLACE1004282, PLACE1004441, PLACE1004450, PLACE1004520, PLACE1004648, PLACE1005003, PLACE1005426, PLACE1006071, PLACE1006073, PLACE1006290, PLACE1007081, PLACE1007845, PLACE1008716, PLACE1008744, PLACE1008985, PLACE1010251, PLACE1010784, PLACE1010968, PLACE1011116, PLACE3000181, PLACE3000213, PLACE4000354, THYRO1000036, THYRO1000196, THYRO1000584, THYRO1000956, THYRO1001266, Y79AA1000270, Y79AA1000426, Y79AA1001727, Y79AA1001795, Y79AA1002022, Y79AA1002213,

[0252] The following 55 clones were categorized into signal transduction-associated proteins.

HEMBA1000006, HEMBA1002195, HEMBA1002227, HEMBA1002551, HEMBA1005084, HEMBA1005929, HEMBA1006658, HEMBA1006916, MAMMA1000881, MAMMA1001150, MAMMA1001310, MAMMA1002142, NT2RM2001902, NT2RP1001020, NT2RP1001031, NT2RP2001469, NT2RP2001529, NT2RP2001769, NT2RP2003179, NT2RP2003545, NT2RP2004670, NT2RP3000011, NT2RP3000022, NT2RP3000172, NT2RP3000201, NT2RP3000820, NT2RP3003527, NT2RP3003849, NT2RP3003874, NT2RP3004067, NT2RP4000634, NT2RP4000962, OVARC1000255, OVARC1000529, OVARC1000916, OVARC1001338, OVARC1001569, PLACE1002329, PLACE1003135, PLACE1003598, PLACE1005519, PLACE1006208, PLACE1008282, PLACE1008297, PLACE1010081, PLACE1011364, PLACE1011824, THYRO1001457, THYRO1001593, THYRO1001700, THYRO1001770, Y79AA1000777, Y79AA1000967, Y79AA1002376, Y79AA1002381,

[0253] The following 80 clones were categorized into transcription -associated proteins.

HEMBA1000462, HEMBA1000671, HEMBA1001297, HEMBA1001390, HEMBA1001886, HEMBA1002048, HEMBA1003120, HEMBA1003497, HEMBA1004785, HEMBA1005230, HEMBA1005246, HEMBA1006276, HEMBA1006572, HEMBA1007226, HEMBB1000106, HEMBB1000905, HEMBB1001959, HEMBB1002051, HEMBB1002661, MAMMA1001094, MAMMA1001532, MAMMA1001615, NT2RM1000789, NT2RM2000632, NT2RM2000773, NT2RM4000326, NT2RP1000271, NT2RP1000468, NT2RP2000092, NT2RP2000610, NT2RP2000712, NT2RP2000739, NT2RP2001538, NT2RP2001662, NT2RP2001817, NT2RP2001948, NT2RP2002564, NT2RP2002974, NT2RP2003138, NT2RP2003302, NT2RP2003940, NT2RP2004108, NT2RP2004847, NT2RP2005247, NT2RP2005391, NT2RP2005535, NT2RP2005774, NT2RP2005941, NT2RP2006092, NT2RP3000148, NT2RP3000232, NT2RP3000378, NT2RP3000652, NT2RP3001976, NT2RP3004090, NT2RP3004119, NT2RP3004294, OVARC1001049, OVARC1001086, OVARC1001132, OVARC1001807, PLACE1000258, PLACE1000442, PLACE1000907, PLACE1003529, PLACE1004166,

Y79AA1000967//GGCACAGACACCATCCTTGA//SEQ ID NO:4461
Y79AA1001056//ACAAATGAGCCTGAAAAGTC//SEQ ID NO:4462
Y79AA1001062//TGGTCCTCACTGCCTTCAAA//SEQ ID NO:4463
Y79AA1001090//AGTGCCCTCAAAGCTCCAGT//SEQ ID NO:4464
Y79AA1001212//ACGAAAGCACTCAAATGTCA//SEQ ID NO:4465
Y79AA1001272//GAATGAAATGTGGTTGAGCA//SEQ ID NO:4466
Y79AA1001426//AATGATTCGGGGCAGCAGGA//SEQ ID NO:4467
Y79AA1001427//GAGAGAGACACACACAGAAA//SEQ ID NO:4468
Y79AA1001523//AGTTTTATACCAGCATTGGC//SEQ ID NO:4469
Y79AA1001530//GGTGTAGAAGTAAAATGGGA//SEQ ID NO:4470
Y79AA1001592//GATTGTGTTCTCTTACTCCT//SEQ ID NO:4471
Y79AA1001727//GCTCCACCTGACGTTCTTTA//SEQ ID NO:4472
Y79AA1001795//GTCTCCCATATCGCTGTCTT//SEQ ID NO:4473
Y79AA1001803//CACTTTCTAATAACCCTTGG//SEQ ID NO:4474
Y79AA1001863//TTGGGATTGGAAACCGCATT//SEQ ID NO:4475
Y79AA1001874//AGAAACCACTGAGGCCCAAG//SEQ ID NO:4476
Y79AA1002058//CAGAAGCAGAAGCAGGAGCA//SEQ ID NO:4477
Y79AA1002121//ATTTACTGCGATTTCTCCTG//SEQ ID NO:4478
Y79AA1002129//GAGTTTCTTTGCTAGTTCCA//SEQ ID NO:4479
Y79AA1002334//ATATTTGTGTTTGCCTTGGG//SEQ ID NO:4480
Y79AA1002373//GGATGGCTGGGTCAAATGGT//SEQ ID NO:4481
Y79AA1002376//AATGATGGCTAGGGTGACTT//SEQ ID NO:4482
Y79AA1002378//TCTTCCCACATTCGTTACAC//SEQ ID NO:4483
Y79AA1002381//AGGGAGTAGATGTTGGTAAA//SEQ ID NO:4484

[0279] The result of expression frequency analysis is shown in Table 367. Only clones with correlation coefficient of 0.9 or higher are indicated in this Table. Clones that are not presented in the Table include clones for which the assay could not performed because of low expression levels thereof in internal standard samples or because of unexpectedly smaller or larger sizes of the PCR products.

[0280] Among the clones that could be analyzed, clones of which expression levels increased by two fold in response to the IL-1. stimulation 1 or 7 hours after the stimulation are: NT2RM2000514, NT2RP3001159, MAMMA1001237 and MAMMA1000614.

[0281] Clones of which expression levels increased by two fold in response to the TNF-stimulation 1, 3 or 7 hours after the stimulation are:
NT2RM2000582, NT2RM2002109, NT2RP1000679, NT2RP2003664, NT2RP2005597, NT2RP2004108, NT2RP3001592, NT2RP3002738, NT2RP3004133, NT2RP3004321, NT2RP3004557, NT2RP3004294, MAMMA1001237, MAMMA1000141, MAMMA1000788, MAMMA1002070, PLACE1002547, PLACE1003573, PLACE1004305, PLACE1008744, PLACE1011181, PLACE1010713, PLACE1010011, Y79AA1000776, Y79AA1002129,

[0282] Among the clones of which expression levels increased in response to IL-1. stimulation, MAMMA1001237 was a clone of which expression level was varied in response to TNF-. stimulation. Among clones showing higher expression levels (with relative value of 5 or higher) prior to the stimulation, PLACE1002080 is an example of clones

of which expression was suppressed by the stimulation. The expression of the clone decreased by three or more fold in response to the stimulation. These genes were found to be associated with inflammatory reaction induced by IL-1. or TNF-..

[0283] In Example 15, the genes of which expression levels were varied by culturing in the presence of TNF-. were analyzed by hybridization with high-density DNA filter. As for 3 clones (NT2RP3004557, NT2RP3004294 and PLACE 1002547), the results obtained by ATAC-PCR method were similar to those obtained by hybridization method. However, the results obtained by ATAC-PCR method were not necessarily consistent with those obtained by the hybridization method. Possible reasons for the inconsistency are the difference in cells used between the two experiments, unavailability of some data in the ATAC-PCR experiment, and the difference in the method of data treatment.

Table 28

Expression of each cDNA in human tissues (The Table also contains clones with no description in Examples)

Clone_name	Heart	Lung	P.gland	Thymus	Brain	Kidney	Liver	Spleen
GAPDH(Cr1)	38.210	32.670	23.820	13.580	11.230	21.120	24.910	22.440
βactin(Cr2)	279.280	368.870	111.100	117.500	92.880	114.650	82.990	256.790
ADRGL1000005	53.882	23.005	32.749	22.858	26.564	24.940	22.644	27.001
ADRGL1000007	94.778	85.185	160.457	67.191	101.768	62.489	67.150	73.543
ADRGL1000009	11.141	50.520	10.357	7.177	6.013	5.219	14.272	21.225
ADRGL1000011	71.656	24.579	29.358	19.473	24.898	30.747	49.220	22.221
ADRGL1000027	36.238	25.252	20.855	7.328	11.196	14.298	19.658	11.288
ADRGL1000058	66.209	129.497	55.226	49.241	30.219	55.872	67.027	243.436
ADRGL1000069	38.630	23.459	28.991	12.540	27.353	33.633	28.774	20.911
ADRGL1000077	97.465	63.656	448.427	83.412	71.108	53.740	67.906	89.439
ADRGL1000092	89.423	45.692	55.810	26.033	44.148	73.339	96.037	73.091
ADRGL1000099	73.675	24.424	36.128	17.024	25.964	41.391	42.837	29.666
ADRGL1000136	141.745	63.974	77.017	24.777	33.549	58.986	295.009	84.985
ADRGL1000147	394.563	155.829	271.210	92.899	165.627	251.266	253.420	150.294
ADRGL1000159	50.073	25.425	39.296	15.194	16.125	20.040	33.720	23.278
ADRGL1000160	69.386	31.051	59.416	20.154	39.799	27.027	47.169	20.716
ADRGL1000171	57.047	23.011	43.063	23.860	40.581	59.814	117.055	32.630
ADRGL1000181	45.892	18.666	34.476	15.434	34.225	32.962	39.693	16.334
BGGI11000015	153.242	42.337	92.865	41.003	45.168	88.524	85.990	73.392
BGGI11000016	177.367	94.731	119.688	34.159	30.249	98.806	98.783	39.204
BGGI11000017	84.712	32.614	38.131	20.878	18.769	32.340	39.666	20.750
BGGI11000022	52.468	20.452	67.167	12.167	11.158	18.241	19.197	11.937
BGGI11000031	30.008	17.072	40.883	12.585	13.313	15.525	16.757	13.406
BGGI11000042	49.926	36.336	51.176	26.964	43.122	43.770	49.107	38.776
BGGI11000046	31.618	26.472	34.182	31.854	12.650	25.784	18.430	25.385
BNGH41000020	5031.103	2993.496	1444.841	537.162	6973.542	6029.124	3350.527.	3649.144
BNGH41000025	91.717	35.026	73.901	27.713	30.765	36.523	37.596:	47.074
BNGH41000026	176.757	77.439	98.345	35.807	56.991	91.310	75.797	70.241
BNGH41000027	65.029	56.353	25.896	22.494	12.763	23.748	17.836	23.859
BNGH41000035	148.779	66.776	119.727	56.576	60.996	96.959	72.461	64.458
BNGH41000037	79.500	29.611	43.438	18.317	20.857	36.272	27.525	24.771
BNGH41000042	224.484	110.084	168.448	104.351	102.259	125.323	86.783	122.959
BNGH41000048	56.144	32.253	54.063	14.729	27.312	22.435	29.566	28.937
BNGH41000056	67.258	18.694	30.075	15.602	10.072	20.735	16.100	7.642
BNGH41000087	98.262	46.173	77.657	35.329	40.900	50.029	50.841	45.285
BNGH41000091	50.895	16.985	28.392	10.147	5.469	22.794	10.725	12.410
BNGH41000157	69.043	34.730	40.597	18.088	27.072	22.074	25.410	24.950
BNGH41000169	44.850	21.770	28.655	11.403	25.991	28.509	25.634	25.843
BNGH41000181	17.163	15.689	13.948	3.996	9.287	13.139	15.553	16.575
BNGH41000198	81.510	36.250	60.860	20.585	26.929	35.751	31.695	28.325
BNGH41000219	30.302	25.156	22.187	13.757	11.208	15.235	27.285	35.709
BNGH41000229	252.790	65.948	93.499	51.108	92.555	101.245	96.716	78.266
BNGH41000237	85.757	46.997	55.170	26.780	33.764	47.456	37.007	39.131
BNGH41000238	17.744	36.938	42.360	14.922	35.749	42.848	39.238	13.241
BNGH41000243	45.446	23.667	44.798	20.875	10.516	23.918	22.443	27.033
BNGH41000270	60.889	18.651	29.618	10.724	15.979	12.351	19.152	22.314
BRAWH1000004	43.673	28.539	7.640	11.388	19.198	14.903	32.353	23.777
BRAWH1000018	59.409	17.941	102.270	17.107	709.078	25.732	24.214	24.767
BRAWH1000021	104.772	29.951	51.142	21.042	1169.154	55.762	66.754	27.969
BRAWH1000027	152.205	47.310	67.089	32.199	64.521	70.731	79.670	40.928
BRAWH1000029	106.376	49.221	55.840	40.856	59.552	56.487	64.886	100.132
BRAWH1000040	29.419	16.761	31.101	16.622	30.633	18.200	17.998	15.196
BRAWH1000050	161.264	71.786	118.976	51.863	61.542	97.720	81.271	69.194
BRAWH1000051	74.067	34.341	44.047	20.726	30.434	42.055	53.856	24.624
BRAWH1000060	68.789	22.598	35.012	16.493	19.127	38.662	34.923	28.094
BRAWH1000075	17.318	16.898	36.437	8.901	18.133	17.219	9.321	11.200
BRAWH1000081	43.025	12.998	28.267	7.655	123.677	17.673	15.924	9.844
BRAWH1000084	174.384	42.178	80.534	47.752	152.188	77.111	110.167	102.296
BRAWH1000095	118.239	59.676	64.528	28.174	116.975	53.814	746.700	35.985
BRAWH1000096	146.112	44.967	85.882	27.491	145.013	52.880	52.427	58.678
BRAWH1000097	95.841	72.506	174.954	65.637	64.200	73.707	63.827	63.762
BRAWH1000100	11.943	19.037	18.950	13.536	92.145	16.582	16.646	10.218
BRAWH1000101	134.838	57.232	106.632	40.741	96.396	71.642	88.432	57.336

Table 104

NT2RP2003099	69.980	61.964	197.831	28.962	29.485	52.756	36.145	46.753
NT2RP2003108	22.037	23.450	29.734	12.784	12.243	25.414	19.582	14.441
NT2RP2003115	175.202	76.490	219.003	26.090	53.025	89.403	96.086	53.165
NT2RP2003117	132.572	135.106	428.449	65.631	66.802	77.649	41.504	75.169
NT2RP2003121	77.521	49.860	42.009	15.143	26.745	31.652	32.041	27.916
NT2RP2003125	35.377	29.656	27.135	9.957	16.383	12.805	20.265	8.252
NT2RP2003127	29.566	16.867	20.397	5.212	10.531	18.240	19.752	7.540
NT2RP2003129	50.461	54.112	157.477	25.025	29.892	16.686	23.103	33.770
NT2RP2003137	8.001	18.759	14.140	10.321	7.469	15.281	5.429	3.225
NT2RP2003138	52.296	44.278	85.267	21.446	22.368	30.612	24.709	34.031
NT2RP2003146	55.329	37.398	52.403	14.492	12.222	29.608	23.329	32.663
NT2RP2003148	150.386	104.523	330.270	60.524	70.523	90.836	76.602	100.291
NT2RP2003150	26.432	11.157	23.761	15.678	11.132	36.468	7.133	18.954
NT2RP2003157	58.172	46.518	64.963	42.288	23.422	50.314	42.129	48.145
NT2RP2003158	44.248	20.906	37.740	8.136	17.954	27.119	19.062	38.471
NT2RP2003161	19.274	11.968	16.062	2.701	7.578	17.086	7.441	31.024
NT2RP2003164	49.401	19.110	28.830	12.219	12.819	22.155	19.787	34.090
NT2RP2003165	89.985	65.955	218.487	37.132	35.205	34.406	24.887	33.303
NT2RP2003177	43.596	22.142	51.196	11.148	3.934	15.303	13.349	69.154
NT2RP2003179	69.718	46.328	169.618	30.883	22.456	37.444	43.967	45.776
NT2RP2003194	144.137	17.980	22.293	13.420	10.852	20.144	19.065	43.611
NT2RP2003206	7.840	5.369	10.850	6.014	4.029	11.290	7.725	3.709
NT2RP2003210	51.322	21.586	38.521	12.974	17.884	37.608	30.477	29.805
NT2RP2003227	42.906	18.716	24.162	17.143	9.513	37.425	15.949	23.165
NT2RP2003228	58.612	29.572	62.903	22.926	28.577	30.449	37.367	63.378
NT2RP2003230	5.885	10.431	148.181	5.253	9.252	9.617	6.228	22.492
NT2RP2003231	69.197	41.691	59.459	34.789	15.272	58.827	33.617	37.859
NT2RP2003237	30.563	38.860	123.572	28.832	11.050	15.189	9.580	23.097
NT2RP2003239	33.469	21.053	50.845	20.348	11.513	25.692	7.484	35.924
NT2RP2003243	145.467	34.182	76.360	17.705	28.702	66.482	55.093	28.921
NT2RP2003265	29.516	23.976	32.673	9.710	15.918	17.608	20.157	14.165
NT2RP2003267	65.087	29.515	67.969	24.282	21.518	34.797	27.241	43.679
NT2RP2003272	41.457	22.351	19.055	27.076	19.762	28.028	26.982	45.977
NT2RP2003277	107.913	82.634	92.986	31.633	32.424	67.812	26.460	53.116
NT2RP2003280	19.151	14.918	20.689	11.633	7.567	43.338	5.070	12.961
NT2RP2003286	21.848	17.740	29.829	11.104	6.965	28.110	26.734	26.233
NT2RP2003293	94.719	83.407	364.260	76.134	56.105	78.539	44.376	97.047
NT2RP2003295	17.874	16.886	18.717	18.256	19.625	15.088	25.617	16.166
NT2RP2003297	9.592	10.816	15.547	2.211	5.615	8.461	10.162	5.662
NT2RP2003300	15.144	16.953	26.519	10.354	14.045	6.847	8.974	11.058
NT2RP2003302	22.071	15.550	64.230	26.397	10.289	12.880	11.722	68.523
NT2RP2003307	22.086	9.418	17.120	5.220	6.112	15.691	17.396	7.096
NT2RP2003308	17.436	24.315	20.930	11.886	7.814	20.422	12.860	31.766
NT2RP2003311	22.001	9.144	13.842	5.360	10.074	18.616	5.176	21.146
NT2RP2003329	44.872	14.471	19.961	10.976	13.401	22.292	12.093	14.770
NT2RP2003339	20.422	19.625	85.412	16.458	12.443	17.818	9.125	13.152
NT2RP2003345	23.118	8.297	17.237	4.695	8.379	12.952	12.259	23.215
NT2RP2003347	12.389	4.636	9.822	7.720	7.500	12.461	7.182	16.011
NT2RP2003367	10.794	19.368	21.160	7.884	14.120	12.142	14.419	13.409
NT2RP2003369	41.141	18.327	38.318	11.072	14.356	33.971	28.126	19.613
NT2RP2003383	55.891	32.218	76.058	21.558	27.536	76.861	50.564	36.175
NT2RP2003390	73.620	57.765	91.034	41.124	35.539	63.744	46.234	42.766
NT2RP2003391	241.564	161.239	277.051	75.828	95.432	220.668	152.546	143.981
NT2RP2003393	11.758	13.507	20.112	4.687	11.809	12.940	19.991	21.749
NT2RP2003394	7.323	9.816	9.506	2.871	10.713	1.307	6.346	14.753
NT2RP2003401	25.259	3.933	8.376	2.832	4.096	7.246	16.159	7.442
NT2RP2003403	31.239	26.205	109.072	18.680	14.206	9.380	14.946	8.745
NT2RP2003433	79.603	33.403	70.460	19.431	29.526	42.730	34.783	28.629
NT2RP2003445	38.525	33.248	95.090	23.648	21.333	27.951	21.347	33.662
NT2RP2003446	67.228	39.971	49.302	18.878	21.829	54.339	39.113	29.464
NT2RP2003456	1.902	13.833	10.178	7.437	1.522	5.049	1.410	3.486
NT2RP2003466	72.001	27.022	47.862	12.506	26.814	66.543	51.004	41.515
NT2RP2003469	35.915	29.791	90.766	19.568	17.254	24.857	16.952	39.575
NT2RP2003470	20.820	31.916	84.744	64.680	20.126	61.522	22.215	98.657

[0284] Expression of each cDNA in human tissues (The Table also contains clones without description in Examples)

Table 185

Expression of each cDNA in human pulmonary arterial endothelial cells cultured in a medium containing bovine serum albumin, glycated bovine serum albumin or advanced glycation endproduct of bovine serum albumin (This table also contains clones without description in Examples).

In the table, EC_G_B/EC_BSA and EC_A_B/EC_BSA represent the ratios EC_glycated_BSA/EC_BSA and EC_AGE_BSA/EC_BSA, respectively.

Clone_name	EC_BSA	EC_glycated_BSA	EC_AGE_BSA	EC_G_B /EC_BSA	EC_A_B /EC_BSA
GAPDH(Cr1)	100.81	134.21	115.16	1.33	1.14
βactin(Cr2)	1101.9	1092.57	997.36	0.99	0.91
ADRGL1000005	26.88	38.27	36.13	1	1
ADRGL1000007	117.89	127.25	133.21	1.08	1.13
ADRGL1000009	29.18	25.65	26.05	1	1
ADRGL1000011	88.9	117.33	142.9	1.32	1.61
ADRGL1000027	33.24	40.53	43.02	1.01	1.08
ADRGL1000058	153.41	208.84	180.05	1.36	1.17
ADRGL1000069	16.8	21.77	29.81	1	1
ADRGL1000077	25.74	24.72	32.86	1	1
ADRGL1000092	84.52	84.15	121.76	1	1.44
ADRGL1000099	76.19	91.53	106.01	1.2	1.39
ADRGL1000136	52.34	44.76	63.06	0.86	1.2
ADRGL1000147	46.08	45.18	52.15	0.98	1.13
ADRGL1000159	31.52	40.24	42.72	1.01	1.07
ADRGL1000160	52.34	60.37	62.29	1.15	1.19
ADRGL1000171	21.46	16.78	25.59	1	1
ADRGL1000181	37.44	45.71	43.65	1.14	1.09
BGGI11000015	52.42	71	65.47	1.35	1.25
BGGI11000016	127.44	122.93	147.57	0.96	1.16
BGGI11000017	25.65	25.74	31.33	1	1
BGGI11000022	32.82	35.19	25.56	1	1
BGGI11000031	44.42	43.8	40.25	0.99	0.91
BGGI11000042	120.38	146.44	165.42	1.22	1.37
BGGI11000046	74.72	58.85	84.95	0.79	1.14
BNGH41000020	4286.08	3584.67	4330.96	0.84	1.01
BNGH41000025	216.67	223.74	257.06	1.03	1.19
BNGH41000026	25.76	28.16	35.52	1	1
BNGH41000027	29.23	23.83	17.86	1	1
BNGH41000035	280.32	238.34	305.66	0.85	1.09
BNGH41000037	59.14	54.86	54.58	0.93	0.92
BNGH41000042	356.1	324.08	411.07	0.91	1.15
BNGH41000048	1201.37	869.03	739.91	0.72	0.62
BNGH41000056	33.94	31.4	40.01	1	1
BNGH41000087	77.58	81.76	91.07	1.05	1.17
BNGH41000091	21.05	21.23	26.82	1	1
BNGH41000157	81.11	57.28	77.46	0.71	0.95
BNGH41000169	21.1	17.59	22.53	1	1
BNGH41000181	63.54	56.92	70.08	0.9	1.1
BNGH41000198	32.53	26.38	34.37	1	1

NT2RP2003157	374.37	442.24	400.41	1.18	1.07
NT2RP2003158	231.08	288.73	253.69	1.25	1.1
NT2RP2003161	34.77	26.93	44.88	1	1.12
NT2RP2003164	18.05	15.52	15.05	1	1
NT2RP2003165	69.08	79.91	55.2	1.16	0.8
NT2RP2003177	37.26	42.87	32.54	1.07	1
NT2RP2003179	29.18	29.09	29.31	1	1
NT2RP2003194	59.15	54.01	59.89	0.91	1.01
NT2RP2003206	19.76	17.37	20.74	1	1
NT2RP2003210	55.61	57.1	56.95	1.03	1.02
NT2RP2003227	44.26	36.25	42.02	0.9	0.95
NT2RP2003228	316.03	318.05	254.76	1.01	0.81
NT2RP2003230	60.62	70.36	78.27	1.16	1.29
NT2RP2003231	32.24	45.25	30.88	1.13	1
NT2RP2003237	58.21	57.82	43.28	0.99	0.74
NT2RP2003239	38.81	42.21	33.65	1.06	1
NT2RP2003243	31.38	28.21	20.39	1	1
NT2RP2003265	32.54	26.78	29.27	1	1
NT2RP2003267	22.93	16.8	22.9	1	1
NT2RP2003272	68.86	81.67	90.05	1.19	1.31
NT2RP2003277	71.96	63.84	62.77	0.89	0.87
NT2RP2003280	52.54	73.49	53.99	1.4	1.03
NT2RP2003286	28.04	26.78	18.22	1	1
NT2RP2003293	71.28	70.24	67.8	0.99	0.95
NT2RP2003295	36.22	40.22	38	1.01	1
NT2RP2003297	28.34	44.52	29.01	1.11	1
NT2RP2003300	100.95	194.28	183.07	1.92	1.81
NT2RP2003302	32.43	34.18	33.1	1	1
NT2RP2003307	14.7	15.78	14.43	1	1
NT2RP2003308	19.05	28.68	21.62	1	1
NT2RP2003311	44	44.95	40.84	1.02	0.93
NT2RP2003329	30.35	27.96	36.19	1	1
NT2RP2003339	54.66	75.38	65.37	1.38	1.2
NT2RP2003345	26.69	20.05	33.5	1	1
NT2RP2003347	20.75	25.26	28.9	1	1
NT2RP2003367	22.37	24.09	19.78	1	1
NT2RP2003369	14.09	18.3	15.2	1	1
NT2RP2003383	44.77	49.15	38.88	1.1	0.89
NT2RP2003390	74.89	87.87	85.57	1.17	1.14
NT2RP2003391	78.99	105.52	90.86	1.34	1.15
NT2RP2003393	24.3	31.69	32.88	1	1
NT2RP2003394	61.47	72.93	77.78	1.19	1.27
NT2RP2003401	29.3	36.77	36.33	1	1
NT2RP2003403	29.08	33.77	22	1	1
NT2RP2003433	126.57	190.47	111.81	1.5	0.88
NT2RP2003445	35.12	50.9	41.57	1.27	1.04
NT2RP2003446	36.9	48.42	57.41	1.21	1.44
NT2RP2003456	31.7	32.64	28.61	1	1
NT2RP2003466	78.87	117.72	134.62	1.49	1.71
NT2RP2003469	38.03	54.3	59.81	1.36	1.5
NT2RP2003470	35.93	52.34	56.11	1.31	1.4
NT2RP2003471	20.57	26.71	27.17	1	1

Y79AA1002373	43.96	55.06	28.34	1.25	0.91
Y79AA1002376	3080.78	3824.05	4481.1	1.24	1.45
Y79AA1002378	73.33	93.61	68.22	1.28	0.93
Y79AA1002381	248.36	288.51	304.13	1.16	1.22
Y79AA1002388	118.82	135.82	129.37	1.14	1.09
Y79AA1002399	36.12	30.1	32.87	1	1
Y79AA1002407	57.84	42.82	52.54	0.74	0.91
Y79AA1002413	78.77	81.36	87.31	1.03	1.11
Y79AA1002416	34.3	30.2	51.99	1	1.3
Y79AA1002429	67.91	69.81	80.19	1.03	1.18
Y79AA1002431	24.66	21.16	23.98	1	1
Y79AA1002433	27.12	18.11	23.63	1	1
Y79AA1002445	78.66	54.58	73.75	0.69	0.94
Y79AA1002461	29.04	24.84	32	1	1
Y79AA1002466	882.69	904.65	782.53	1.02	0.89
Y79AA1002471	53.74	51.26	68.91	0.95	1.28
Y79AA1002472	121.95	127.4	127.11	1.04	1.04
Y79AA1002474	53.33	40.85	47.18	0.77	0.88
Y79AA1002482	103.36	111.11	116.07	1.07	1.12
Y79AA1002487	30.92	25.8	32.51	1	1
Y79AA1002490	101.4	90.92	90.54	0.9	0.89
Y79AA1002493	107.88	125.54	105.75	1.16	0.98
ZRV6C1006278	46.63	30.08	32.23	0.86	0.86

Table 186

Expression of each cDNA in undifferentiated NT2 cells, in NT2 cells cultured in the presence of retinoic acid, or in NT2 cells that were cultured in the presence of retinoic acid and then further cultured in the presence of cell-division inhibitor added (This table also contains clones without description in Examples)

In the table, NT2, NT2_RA, and NT2_RA_INHIB represent untreated NT2 cells, retinoic acid-treated NT2 cells, and retinoic acid/inhibitor-treated NT2 cells, respectively. The assay was performed in triplicate (n=3), and each result was shown in the column of exp.1, exp.2, or exp.3. In addition, "t-test N/R" and "t-test N/I" represent results of test for significance of difference between the untreated cells and the retinoic acid-treated cells, and between the untreated cells and the retinoic acid/inhibitor-treated cells, respectively. The results of the test

are shown in the columns of *:p<0.05 and **:p<0.01.

Clone	NT2			NT2 RA			NT2 RA INHIB			ttest N/R	+ −	ttest N/I	+ −
	exp.1	exp.2	exp.3	exp.1	exp. 2	exp.3	exp.1	exp.2	exp.3				
GAPDH(Cr1)	3.53	1.08	0.98	2.92	2.49	2.8	1.76	2.59	1.52				
β actin(Cr2)	155.4	118	99.68	148.5	110.7	101.3	114.7	105.8	151.1				
ADRGL1000005	4.01	2.03	1.55	4.05	3.65	3.6	2.27	2.93	4.24				
ADRGL1000007	11.08	5.73	7.92	15.42	10.6	13.87	8.99	8.17	9.15				
ADRGL1000009	1.11	0.72	1.04	1.66	1.89	1.03	1.22	1.62	1.58			*	+
ADRGL1000011	4.27	2.7	2.85	4.32	4.35	3.38	2.76	3.27	3.06				
ADRGL1000027	1.83	0.38	0.56	0.97	0.62	0.99	0.92	1.33	1.5				
ADRGL1000058	3.65	2.58	1.37	2.92	3.36	2.75	2.25	3.51	2.7				
ADRGL1000069	3.25	1.85	3.28	1.86	2.53	2.85	2.01	2.89	2.7				
ADRGL1000077	13.48	10.41	6.71	19.62	17.92	22.59	11.6	16.66	19.34	*	+		
ADRGL1000092	5.73	2.8	4.51	7.31	5.01	4.83	3.24	6.16	7.22				
ADRGL1000099	5.64	3.42	2.08	5.59	3.73	4.24	3.98	3.98	4.06				
ADRGL1000136	9.97	3.52	4.19	5.77	4.73	5.86	6.61	5.16	5.49				
ADRGL1000147	23.09	13.85	11.7	14.77	14.96	14.89	17.7	13.3	19.47				
ADRGL1000159	6.11	2.22	3.37	5.24	2.88	4.15	2.76	2.93	3.59				
ADRGL1000160	7.16	3.48	4.19	5.94	4.59	3.41	3.95	4.67	4.25				
ADRGL1000171	4.84	2.99	3.23	3.52	4.19	4.37	2.55	3.88	3.45				
ADRGL1000181	5.1	3.65	2.6	3.16	4.06	2.97	2.64	3.06	3.44				
BGGI11000015	13.95	6.83	6.72	9.61	9.19	10.24	9.94	10.66	10.13				
BGGI11000016	15.49	5.92	7.09	11.88	11.38	8.72	11.82	10.98	10.51				
BGGI11000017	7.89	2.99	3.25	4.94	4.94	4.93	3.55	4.27	3.52				
BGGI11000022	8.77	5.14	5.91	7.12	7.05	4.54	5.71	5.59	5.9				
BGGI11000031	4.71	2.16	2.74	4.09	3.29	3.96	4.02	3.67	2.33				
BGGI11000042	6.37	5.24	3.74	5.63	6.22	4.36	4.66	5.2	4.04				
BGGI11000046	19.01	12.57	9.23	12.39	15.7	12.37	8.8	10.92	9.17				
BNGH41000020	859	910.1	603	164	319.2	267.4	638.2	771.6	845.4	**	-		
BNGH41000025	5.35	2.06	2.09	2.76	2.76	3.77	4.23	2.01	3.06				
BNGH41000026	16.2	7.69	7.05	9.34	11.37	9.66	10.13	7.16	10.71				
BNGH41000027	2.31	2.18	2.5	2.9	3.01	2.82	3.68	3.48	4.21	**	+	**	+
BNGH41000035	14.57	8.83	9.36	10.92	9.55	14.75	15.02	15.18	12.2				
BNGH41000037	10.56	7.46	6.2	8.16	9.21	6.42	3.37	5.45	4.98				
BNGH41000042	77.1	50.85	58.45	47.64	53.39	62.67	28.12	35.48	23.44			*	-
BNGH41000048	3.5	2.19	1.91	4.28	2.87	2.4	1.63	3.01	1.78				
BNGH41000056	2.57	2.01	1	1.91	2.63	2.15	1.41	2.4	1.79				
BNGH41000087	9.84	5.84	5.53	12.49	10.24	10.25	11.74	9.68	8.53				
BNGH41000091	3.37	2.59	1.21	3.29	3.01	1.55	2.95	2.57	2.13				
BNGH41000157	10.63	5.64	6.15	8.53	9.05	7.74	6.38	6.68	5.75				
BNGH41000169	3.77	4.34	3.82	4.9	3.48	3.32	3.4	4.16	4.19				
BNGH41000181	2.47	1.59	1.84	2.93	2.1	1.8	1.7	2.66	1.59				
BNGH41000198	8.13	4.64	3.79	5.48	4.35	5.59	4.3	4.15	4.35				
BNGH41000219	9.61	3.92	4.87	4.17	5.29	5.45	5.24	7.12	7.13				
BNGH41000229	19.61	13.28	8.68	10.86	11.27	9.36	7.9	9.5	10.85				
BNGH41000237	10.9	5.47	6.45	6.65	6.97	7.79	6.36	6.25	5.44				
BNGH41000238	4.58	7	3.45	5.91	4.68	4.34	4.33	5.44	4.22				
BNGH41000243	13.85	8.69	8.48	10.19	9.71	8.97	8.23	4.87	5.54				
BNGH41000270	5.83	2.62	2.35	2.3	3.05	3.44	2.59	3.49	1.3				
BRAWH1000004	4.19	2.83	2.48	5.04	3.15	3.26	1.44	3.45	2.05				
BRAWH1000018	4.85	1.95	2.29	7.47	8.8	8.85	8.68	6.61	7.96	**	+	*	+
BRAWH1000021	6.52	5.06	5.87	5.09	6.94	6.44	2.89	6.23	4.28				
BRAWH1000027	11.64	8.86	7.19	8.24	10.39	11.51	5.58	7.13	8.24				
BRAWH1000029	9.58	5.15	3.52	6.01	6.72	6	5.08	5.12	5.84				
BRAWH1000040	4.6	1.89	2.14	2.92	2.71	2.7	2.92	2.5	3.01				
BRAWH1000050	11.48	4.95	5.19	9.74	7.25	8.62	8.25	8.09	8.93				
BRAWH1000051	8.18	3.93	3.19	6.15	5.72	6.02	5.01	4.25	4.44				

Table 273

NT2RP2003129	3.68	2.64	1.93	5.72	5.89	5.75	3.03	4.40	2.82	**	+		
NT2RP2003137	2.40	2.79	2.71	6.74	6.38	5.76	4.22	6.41	4.31	**	+	*	+
NT2RP2003138	6.42	2.67	2.97	5.99	6.92	3.98	5.12	3.06	1.92				
NT2RP2003146	4.44	2.51	1.78	3.73	3.26	2.77	3.76	2.57	1.66				
NT2RP2003148	9.10	6.45	5.51	11.73	13.86	11.19	8.71	8.13	7.46	*	+		
NT2RP2003150	3.26	2.20	1.35	8.65	2.99	4.86	3.92	2.84	8.35				
NT2RP2003157	7.49	3.86	3.67	8.41	10.43	9.55	4.96	6.45	5.87	*	+		
NT2RP2003158	1.98	1.89	2.17	2.26	3.00	2.46	2.43	2.76	2.85			*	+
NT2RP2003161	1.04	1.33	0.76	2.12	4.38	4.18	1.59	2.84	8.91	*	+		
NT2RP2003164	2.83	1.78	1.70	2.90	2.78	2.57	2.53	2.97	2.44				
NT2RP2003165	4.31	2.10	2.06	5.98	4.84	6.84	5.12	3.81	4.72	*	+		
NT2RP2003177	3.18	2.52	2.22	3.53	2.99	3.63	4.35	2.80	2.79				
NT2RP2003179	4.54	3.39	3.36	5.90	7.70	7.29	4.85	4.79	6.24	**	+		
NT2RP2003194	16.94	9.59	9.74	7.86	8.77	6.84	7.23	6.50	9.93				
NT2RP2003206	0.19	0.73	0.54	2.02	2.10	1.11	1.07	1.15	1.17	*	+	*	+
NT2RP2003210	5.52	2.50	2.65	2.94	4.61	3.60	3.44	3.99	4.15				
NT2RP2003227	2.55	1.52	2.78	3.96	4.66	3.48	2.52	3.60	4.44	*	+		
NT2RP2003228	5.50	4.11	4.96	4.07	4.64	3.51	3.63	3.86	2.66				
NT2RP2003230	1.04	1.41	1.38	3.75	3.72	3.44	8.77	4.96	7.21	**	+	**	+
NT2RP2003231	6.83	5.52	4.87	9.61	7.64	6.47	5.75	5.89	8.09				
NT2RP2003237	4.46	2.56	2.35	5.51	7.13	6.33	3.56	4.31	3.67	*	+		
NT2RP2003239	4.50	2.01	3.71	6.44	6.32	5.76	4.01	4.23	4.42	*	+		
NT2RP2003243	5.46	3.20	3.57	7.44	6.11	7.58	5.91	6.40	3.87	*	+		
NT2RP2003265	5.61	3.24	3.60	7.47	8.92	7.01	5.38	4.10	6.74	*	+		
NT2RP2003267	3.97	3.06	3.71	7.15	8.86	6.88	4.28	4.40	5.84	**	+		
NT2RP2003272	5.37	3.98	5.63	6.49	6.56	6.62	7.54	6.51	7.61	*	+	*	+
NT2RP2003277	9.14	5.91	4.66	7.52	10.35	9.11	9.97	7.77	15.8				
NT2RP2003280	3.01	2.25	1.41	4.02	6.71	7.68	6.13	4.20	7.59	*	+	*	+
NT2RP2003286	3.53	1.84	2.37	2.62	3.15	2.83	2.96	2.70	4.01				
NT2RP2003293	6.85	4.64	6.03	12.22	12.54	11.97	6.66	5.15	8.8	**	+		
NT2RP2003295	4.81	3.25	3.18	3.96	8.36	5.27	4.16	4.98	3				
NT2RP2003297	1.97	1.06	1.42	2.82	3.09	2.49	1.97	1.89	1.68	*	+		
NT2RP2003300	5.99	4.89	4.68	7.75	7.40	7.47	7.28	9.19	9.08	**	+	*	+
NT2RP2003302	4.65	3.24	4.39	8.90	10.20	7.29	4.36	7.27	5.11	**	+		
NT2RP2003307	1.67	1.09	0.57	2.24	1.67	2.40	2.82	1.84	1.76				
NT2RP2003308	3.09	2.17	1.85	4.09	5.19	2.83	3.04	2.74	3.16				
NT2RP2003311	6.85	3.58	2.13	4.65	6.66	4.36	3.88	3.65	4.23				
NT2RP2003329	3.07	1.86	1.87	3.19	5.07	3.49	3.77	3.82	5.96				
NT2RP2003339	2.38	1.55	1.29	2.90	3.98	3.91	2.69	3.47	2.24	*	+		
NT2RP2003345	1.83	1.44	1.40	1.51	1.52	1.92	2.28	2.65	1.28				
NT2RP2003347	1.48	2.10	1.67	2.03	5.75	1.76	2.44	3.10	4.09			*	+
NT2RP2003367	1.26	0.98	1.42	1.39	1.59	1.55	1.21	2.14	1.04				
NT2RP2003369	3.82	2.31	1.37	1.62	2.10	1.87	3.19	2.85	1.99				
NT2RP2003383	7.18	3.57	4.41	16.30	14.96	15.98	8.79	9.62	11.29	**	+	*	+
NT2RP2003390	9.92	6.14	6.73	11.71	12.19	9.52	7.92	9.43	8.34				
NT2RP2003391	35.23	21.64	23.50	36.95	36.23	27.51	23.69	17.29	17.85				
NT2RP2003393	2.40	1.57	1.83	4.13	5.18	3.56	3.96	4.34	3.87	*	+	**	+
NT2RP2003394	4.02	2.41	2.76	12.16	9.99	10.68	6.12	6.15	3.96	**	+		
NT2RP2003401	2.33	1.80	1.86	3.02	4.68	2.41	3.02	4.51	3.57			*	+
NT2RP2003403	1.23	1.40	1.41	3.20	3.23	4.51	3.04	3.80	3.41	**	+	**	+
NT2RP2003433	8.96	4.52	3.52	6.71	5.66	5.39	7.4	6.01	5.01				
NT2RP2003445	3.20	3.09	2.41	6.94	6.16	6.94	13.01	11.43	14.04	**	+	**	+
NT2RP2003446	5.05	4.02	2.72	4.09	6.31	3.82	5.45	4.95	5.35				
NT2RP2003456	4.21	2.96	2.69	10.80	8.14	8.43	6.15	5.44	4.71	**	+	*	+
NT2RP2003466	5.26	3.68	3.82	5.95	5.44	4.60	3.82	5.23	9				
NT2RP2003469	3.53	2.12	2.45	3.89	4.69	5.28	2.75	4.01	3.09	*	+		

HEMBA1003885	4.59	4.82	7.14	9.19	6.32	8.41		
HEMBA1003887	3.58	4.93	7.7	8.65	7.93	8.18		
HEMBA1003890	4.2	4.48	7.18	7.53	9.1	6.26		
HEMBA1003893	4.38	6.39	9.53	8.75	13.24	9.94		
HEMBA1003896	4.15	4.15	10.62	7.4	9.12	6.43		
HEMBA1003902	1.39	3.78	5.09	4.91	6.42	5.1		
HEMBA1003904	0.87	2.16	2.46	2.82	4.32	2.11		
HEMBA1003908	1.18	1.3	2.89	2.12	5.25	1.43		
HEMBA1003926	14.46	12.2	39.79	45.5	34.97	55.56		
HEMBA1003937	2.75	3.31	5.38	4.3	6.85	4.57		
HEMBA1003939	2.43	2.48	6.56	8.3	13.32	8.04		
HEMBA1003940	2.45	3.08	5.01	4.29	6.22	5.55		
HEMBA1003941	1.4	2.26	2.48	3.37	4.57	4.42	*	+
HEMBA1003942	1.63	2.88	3.13	2.01	3.85	2.22		
HEMBA1003945	12.57	13.75	22.75	20.99	14.77	19.74		
HEMBA1003949	1.4	1.9	3.53	3.29	6.22	4.14		
HEMBA1003950	3.46	4.86	6.49	14.69	17.53	13.02	**	+
HEMBA1003953	1.91	1.6	5.14	0.72	3.97	1.44		
HEMBA1003958	5.16	3.6	7.47	7.54	9.45	6.64		
HEMBA1003959	2.42	2.72	5.72	5.5	5.5	9.02		
HEMBA1003960	3.25	5.81	34.7	24.04	26.4	28.28		
HEMBA1003966	9.63	8.28	16.73	16.75	17.67	19.84		
HEMBA1003967	1.75	3.06	3.47	3.48	3.6	3.27		
HEMBA1003968	0.97	2.14	2.55	2.49	4.56	1.82		
HEMBA1003974	634.2	699.64	821.36	986.23	1340.97	1248.21	*	+
HEMBA1003976	1.05	1.84	3.36	1.21	3.27	2.04		
HEMBA1003977	1.48	2.07	1.99	1.41	3.49	2.15		
HEMBA1003978	2.91	3.72	3.54	3.77	6.18	3.53		
HEMBA1003981	9.01	6.77	14.06	12.05	11.49	18.27		
HEMBA1003982	102.64	103.61	302.15	380.08	375.9	466.69	*	+
HEMBA1003985	1.18	1.9	2.43	3.21	3.79	2.18		
HEMBA1003987	3.04	2.23	3.1	2.56	4.34	5.53		
HEMBA1003989	1.62	1.77	4.56	3.79	5.12	3.31		
HEMBA1004000	1.63	2.35	5.05	3.46	5.35	4.18		
HEMBA1004006	2.79	2.88	12.86	16.29	22.13	19.73	*	+
HEMBA1004007	0.7	1.92	5.28	3.03	5.18	4.72		
HEMBA1004010	67.4	61.25	98.24	112.56	96.78	136.86		
HEMBA1004011	0.48	1.74	2.18	2.58	3.29	1.62		
HEMBA1004012	0.79	1.84	2.3	3.11	4.8	3.53	*	+
HEMBA1004015	2.68	4.15	5.38	8.68	10.65	9.21	**	+
HEMBA1004024	1.47	2.73	5.65	5.68	8.26	8.07		
HEMBA1004029	1.93	3.1	3.03	4.6	8.38	13.11		
HEMBA1004038	1.04	1.24	1.55	1.18	3.38	1.59		
HEMBA1004042	0.89	1.42	2.22	1.58	4	2.48		
HEMBA1004045	0.28	0.94	2.42	3.07	3.32	2.53		
HEMBA1004048	4.16	4.16	12.1	19.93	14.84	22.3	*	+
HEMBA1004049	3.56	3.18	4.87	4.92	6.83	5.48		
HEMBA1004051	136.19	118.77	205.49	243.62	283.22	223.29	*	+
HEMBA1004053	5.11	4.64	8.92	25.25	27.24	21	**	+
HEMBA1004055	2.28	3.2	4.24	2.15	5.51	2.86		
HEMBA1004056	3.78	3.07	6.73	5.3	10.99	9.56		
HEMBA1004060	0.86	1	1.7	0.78	3.94	1.65		

HEMBA1004061	4.76	3.94	6.44	7.37	12.64	8.57		
HEMBA1004067	10.12	14.76	90.67	108.89	125.21	128.6	*	+
HEMBA1004071	7.51	7.77	16.52	17.31	12.23	13.37		
HEMBA1004074	0.78	1.93	3.97	4.48	7.06	5.69	*	+
HEMBA1004078	3.87	2.95	5.22	6.52	6.2	6.87	*	+
HEMBA1004085	1.05	1.19	2.83	3.57	4.57	2.45		
HEMBA1004086	3.38	4.95	6	8.92	8.09	6.51	*	+
HEMBA1004097	1.18	1.13	2.97	3.66	3.28	2.97		
HEMBA1004100	3.85	4.81	8.96	6.9	9.64	9.55		
HEMBA1004103	2	2.91	6.25	6.25	7.24	7.38		
HEMBA1004110	3	3.77	5.43	4.18	4.23	5.02		
HEMBA1004111	3.96	7.64	44.2	53.81	60.1	57.3	*	+
HEMBA1004124	7.14	10.51	60.12	83.27	97.96	83.59	*	+
HEMBA1004130	3.12	3.46	10.29	9.45	6.84	8.43		
HEMBA1004131	2.14	2.12	3.06	4.08	3.73	3.21	*	+
HEMBA1004132	0.77	2.22	4.84	3.94	6.31	4.2		
HEMBA1004133	0.69	1.77	2.56	3.28	5.17	3.22		
HEMBA1004138	0.89	1.19	3.05	2.21	4.11	1.83		
HEMBA1004143	7.1	7.48	17.43	18.83	15	17.6		
HEMBA1004146	0.89	2.03	3.01	2.96	4.21	2.69		
HEMBA1004148	1.85	1.57	2.13	2.25	3.38	1.99		
HEMBA1004149	1.54	1.44	2.77	2.83	2.59	3.32		
HEMBA1004150	0.49	1.06	2.15	2.31	1.58	1.08		
HEMBA1004154	2.24	1.64	5.28	6.28	7.07	4.61		
HEMBA1004164	1.84	2.23	5.63	6.89	7.13	5.81		
HEMBA1004168	2.16	2.24	4.69	3.9	5.32	7.84		
HEMBA1004199	1.37	1.92	2.34	3.17	3.66	1.8		
HEMBA1004200	0.84	1.98	3	1.5	4.05	1.78		
HEMBA1004201	4.87	5.68	17.64	26.94	32.17	25.65	*	+
HEMBA1004202	7.7	10.5	9.9	18.08	16.29	15.77	**	+
HEMBA1004203	1.63	2.31	3.66	4.5	5.3	4.44	*	+
HEMBA1004207	1.9	3.24	3.62	5.73	6.23	6.2	**	+
HEMBA1004210	1.13	1.72	2.67	1.95	4.14	1.87		
HEMBA1004225	1.1	2.47	5.23	5.96	7.12	5.4		
HEMBA1004227	2.17	4.44	3.86	5.14	5.71	5.16		
HEMBA1004235	2.68	2.91	3.74	5.79	5.78	4.44	*	+
HEMBA1004237	3	3.31	5.23	5.95	4.67	5.47		
HEMBA1004238	2.06	3.24	5.93	5.84	7.64	6.52		
HEMBA1004241	2.32	3.09	3.87	2.74	3.74	3.35		
HEMBA1004242	8.66	13.05	20.15	26.83	32.28	26.48	*	+
HEMBA1004243	1.8	2.09	3.58	2.8	3.03	2.76		
HEMBA1004246	1.6	2.68	5.65	6.18	6.24	6.15		
HEMBA1004247	0.89	2.73	3.74	3.69	4.23	3.37		
HEMBA1004248	4.01	3.54	3.85	5.91	8.31	7.47	**	+
HEMBA1004250	1.55	2.16	2.87	1.91	5.22	1.47		
HEMBA1004252	3.57	3.27	4.8	4.64	5.79	4.28		
HEMBA1004260	2.56	3.08	6.87	7.32	8.16	7.61		
HEMBA1004264	1.26	2.11	2.59	2.16	2.86	1.37		
HEMBA1004267	5.5	5.81	14.29	14.22	12.19	11.57		
HEMBA1004272	1.75	2.31	3.31	2.26	3.84	2.04		
HEMBA1004274	5.83	8.13	58.69	77.19	87.61	76.22	*	+
HEMBA1004275	1	5.4	3.34	1.49	4.49	2.42		

NT2RP2002979	4.41	4.41	8.12	7.03	8.66	7.6		
NT2RP2002980	6.44	6.44	15.09	15.56	11	17.45		
NT2RP2002986	3.87	3.87	7.6	6.68	7.4	7.39		
NT2RP2002987	3.52	3.52	8.23	11.1	9.18	9.4	*	+
NT2RP2002988	14.96	14.96	22.92	30.07	31.87	31.36	**	+
NT2RP2002993	2.97	2.97	4.18	3.8	3.84	2.84		
NT2RP2003000	4.88	4.88	8.34	6.97	9.62	9.97		
NT2RP2003008	4.85	4.85	5.06	3.34	4.76	4.78		
NT2RP2003020	4.45	4.45	44.26	28.35	46.52	34.33		
NT2RP2003032	1.91	1.91	4.02	5.82	6.48	6.59	**	+
NT2RP2003034	4.21	4.21	13.47	13.16	11.15	16.31		
NT2RP2003042	2.15	2.15	3.81	4.57	3.65	4.92		
NT2RP2003050	2.32	2.32	3.56	2.55	2.17	1.83		
NT2RP2003060	7.27	7.27	15.51	21.53	18.91	17.46	*	+
NT2RP2003073	5.61	5.61	8.73	7.06	10.51	8.17		
NT2RP2003099	5.05	5.05	3.67	3.21	3.73	2.84		
NT2RP2003108	3.6	3.6	4.23	5.29	3.91	6.62		
NT2RP2003115	1.68	1.68	5	7.75	4.69	4.84		
NT2RP2003117	2.71	2.71	5.69	3.6	4.66	4.13		
NT2RP2003121	1.83	1.83	3.47	4.03	2.69	3.33		
NT2RP2003125	4.13	4.13	11.44	15.42	12.55	13.66	*	+
NT2RP2003127	2.36	2.36	3.94	1.53	1.66	1.75		
NT2RP2003129	3.43	3.43	7.09	6.08	6.05	5.42		
NT2RP2003137	4.49	4.49	6.14	7.58	8.4	6.46	*	+
NT2RP2003138	4.66	4.66	20.24	16.55	17.45	16.92		
NT2RP2003146	6.2	6.2	24.78	18.5	23.25	25.96		
NT2RP2003148	3.09	3.09	6.73	3.06	4.6	4.04		
NT2RP2003150	1.45	1.45	5.71	3.98	5.2	4.3		
NT2RP2003157	6.93	6.93	34.27	34.29	31.85	32.84		
NT2RP2003158	6.3	6.3	25.32	26.87	28.69	59.31		
NT2RP2003161	2.73	2.73	3.36	2.51	2.82	6.12		
NT2RP2003164	1.96	1.96	2.1	1.28	1.87	2.46		
NT2RP2003165	2.18	2.18	5.94	3.1	3.69	4.84		
NT2RP2003177	1.63	1.63	4.37	2.79	3.03	4.42		
NT2RP2003179	1.23	1.23	4.98	4.08	3.63	7.96		
NT2RP2003194	4.04	4.04	7.2	5.73	6.29	14.77		
NT2RP2003206	1.59	1.59	4.47	1.64	3.52	1.44		
NT2RP2003210	5.06	5.06	15.15	16.14	12.93	15.9		
NT2RP2003227	1.62	1.62	3.97	2.04	3.66	6.28		
NT2RP2003228	6.57	6.57	29.53	29.56	43.94	44.24		
NT2RP2003230	3.51	3.51	7.91	4.49	8.04	8.46		
NT2RP2003231	2.22	2.22	5.59	2.46	3.23	3.83		
NT2RP2003237	2.52	2.52	4.59	4.59	6.4	6.46	*	+
NT2RP2003239	2.3	2.3	4.46	2.97	4.46	4.05		
NT2RP2003243	2.16	2.16	4.13	2.38	3.28	3.98		
NT2RP2003265	3.93	3.93	5.33	4.22	4.88	4.92		
NT2RP2003267	2.73	2.73	3.15	3.24	4.17	7.42		
NT2RP2003272	6.03	6.03	14.8	16.93	23.85	32.58	*	+
NT2RP2003277	3.85	3.85	11.29	5.53	8.39	6.39		
NT2RP2003280	3.47	3.47	9.38	7.67	7.25	6.09		
NT2RP2003286	2.18	2.18	4.23	4.13	5	9.61		
NT2RP2003293	2.98	2.98	6.9	5.66	7.05	7.94		

Table 370

clone	the name of full-length nucleotide sequences	SEQ ID NO of the full-length nucleotide sequences	SEQ ID NO of the deduced amino acid sequences
HEMBA1000006	C-HEMBA1000006	2547	2548
nnnnnnnnnnnn	C-nnnnnnnnnnnn	nnnn	nnnn
HEMBA1000121	C-HEMBA1000121	2551	2552
HEMBA1000128	C-HEMBA1000128	2553	2554
HEMBA1000275	C-HEMBA1000275	2555	2556
HEMBA1000300	C-HEMBA1000300	2557	
HEMBA1000349	C-HEMBA1000349	2558	2559
HEMBA1000443	C-HEMBA1000443	2560	2561
HEMBA1000590	C-HEMBA1000590	2562	2563
HEMBA1000634	C-HEMBA1000634	2564	2565
HEMBA1000713	C-HEMBA1000713	2566	2567
HEMBA1000745	C-HEMBA1000745	2568	2569
HEMBA1000907	C-HEMBA1000907	2570	2571
HEMBA1000940	C-HEMBA1000940	2572	2573
HEMBA1000962	C-HEMBA1000962	2574	2575
HEMBA1001221	C-HEMBA1001221	2576	2577
HEMBA1001228	C-HEMBA1001228	2578	2579
HEMBA1001297	C-HEMBA1001297	2580	
HEMBA1001390	C-HEMBA1001390	2581	2582
HEMBA1001563	C-HEMBA1001563	2583	
HEMBA1001621	C-HEMBA1001621	2584	2585
nnnnnnnnnnnn	C-nnnnnnnnnnnn	nnnn	nnnn
HEMBA1001878	C-HEMBA1001878	2588	2589
HEMBA1002131	C-HEMBA1002131	2590	2591
HEMBA1002163	C-HEMBA1002163	2592	2593
HEMBA1002164	C-HEMBA1002164	2594	2595
HEMBA1002167	C-HEMBA1002167	2596	2597
HEMBA1002178	C-HEMBA1002178	2598	2599
nnnnnnnnnnnn	C-nnnnnnnnnnnn	nnnn	nnnn

NT2RP2002564	C-NT2RP2002564	2902	2903
NT2RP2002674	C-NT2RP2002674	2904	2905
NT2RP2002721	C-NT2RP2002721	2906	2907
NT2RP2002824	C-NT2RP2002824	2908	2909
NT2RP2002942	C-NT2RP2002942	2910	2911
NT2RP2002974	C-NT2RP2002974	2912	2913
NT2RP2002976	C-NT2RP2002976	2914	2915
NT2RP2003042	C-NT2RP2003042	2916	2917
NT2RP2003179	C-NT2RP2003179	2918	2919
NT2RP2003210	C-NT2RP2003210	2920	2921
NT2RP2003369	C-NT2RP2003369	2922	2923
NT2RP2003383	C-NT2RP2003383	2924	2925
NT2RP2003469	C-NT2RP2003469	2926	2927
NT2RP2003545	C-NT2RP2003545	2928	2929
NT2RP2003593	C-NT2RP2003593	2930	2931
NT2RP2003599	C-NT2RP2003599	2932	2933
NT2RP2003655	C-NT2RP2003655	2934	2935
NT2RP2003931	C-NT2RP2003931	2936	2937
NT2RP2004141	C-NT2RP2004141	2938	2939
NT2RP2004179	C-NT2RP2004179	2940	2941
NT2RP2004205	C-NT2RP2004205	2942	2943
NT2RP2004447	C-NT2RP2004447	2944	2945
NT2RP2004495	C-NT2RP2004495	2946	2947
NT2RP2004524	C-NT2RP2004524	2948	2949
NT2RP2004556	C-NT2RP2004556	2950	2951
NT2RP2004606	C-NT2RP2004606	2952	2953
NT2RP2004648	C-NT2RP2004648	2954	2955
NT2RP2004670	C-NT2RP2004670	2956	2957
NT2RP2004794	C-NT2RP2004794	2958	2959
NT2RP2004837	C-NT2RP2004837	2960	2961
NT2RP2004847	C-NT2RP2004847	2962	2963
nnnnnnnnnnnn	C-nnnnnnnnnnnn	nnnn	nnnn
NT2RP2005027	C-NT2RP2005027	2966	
NT2RP2005163	C-NT2RP2005163	2967	2968
NT2RP2005181	C-NT2RP2005181	2969	2970
NT2RP2005247	C-NT2RP2005247	2971	2972
NT2RP2005425	C-NT2RP2005425	2973	2974
NT2RP2005463	C-NT2RP2005463	2975	2976
NT2RP2005514	C-NT2RP2005514	2977	2978
NT2RP2005541	C-NT2RP2005541	2979	2980
NT2RP2005632	C-NT2RP2005632	2981	2982
NT2RP2005878	C-NT2RP2005878	2983	2984
NT2RP2005883	C-NT2RP2005883	2985	2986
NT2RP2005887	C-NT2RP2005887	2987	2988
NT2RP2005941	C-NT2RP2005941	2989	2990
NT2RP2005994	C-NT2RP2005994	2991	2992
NT2RP2006042	C-NT2RP2006042	2993	2994
NT2RP2006269	C-NT2RP2006269	2995	2996
NT2RP2006512	C-NT2RP2006512	2997	2998
NT2RP3000059	C-NT2RP3000059	2999	3000
NT2RP3000063	C-NT2RP3000063	3001	3002
NT2RP3000125	C-NT2RP3000125	3003	3004

1.0e-16:139:36
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32802

F-NT2RP2002942
NEURAL CELL ADHESION MOLECULE L1 PRECURSOR (N-CAM L1).
5.1e-18:153:30
HOMO SAPIENS (HUMAN).
P32004

F-NT2RP2002974
HOMEOBOX PROTEIN SIX5 (DM LOCUS-ASSOCIATED HOMEODOMAIN PROTEIN HOMOLOG) (FRAGMENT).
3.6e-80:187:84
MUS MUSCULUS (MOUSE).
P70178

F-NT2RP2002976
HYPOTHETICAL 149.7 KD PROTEIN IN IRE1-KSP1 INTERGENIC REGION.
2.8e-18:99:47
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P38800

F-NT2RP2003042
PHOSPHATIDYLCHOLINE-STEROL ACYLTRANSFERASE PRECURSOR (EC 2.3.1.43) (LECITHIN-CHOLESTEROL ACYLTRANSFERASE) (PHOSPHOLIPID-CHOLESTEROL ACYLTRANSFERASE) (FRAGMENT).
1.2e-41:135:57
GALLUS GALLUS (CHICKEN).
P53760

F-NT2RP2003138
5'-TG-3'INTERACTING FACTOR (HOMEOBOX PROTEIN TGIF).
3.3e-09:104:45
MUS MUSCULUS (MOUSE).
P70284

F-NT2RP2003179
CARBON CATABOLITE DEREPRESSING PROTEIN KINASE (EC 2.7.1.-).
7.2e-15:96:40
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P06782

F-NT2RP2003210
LONG-CHAIN FATTY ACID TRANSPORT PROTEIN (FATP).
6.2e-69:235:57
MUS MUSCULUS (MOUSE).
Q60714

F-NT2RP2003302
ZINC FINGER PROTEIN 136.
9.7e-52:140:52
HOMO SAPIENS (HUMAN).
P52737

F-NT2RP2003369
SALIVARY PROLINE-RICH PROTEIN PO (ALLELE M) [CONTAINS: PEPTIDE P-D] (FRAGMENT).
0.00020:87:32
HOMO SAPIENS (HUMAN).

C-NT2RP2002824//ENDOSOMAL P24A PROTEIN PRECURSOR (70 KD ENDOMEMBRANE PROTEIN) (PHEROMONE ALPHA-FACTOR TRANSPORTER) (ACIDIC 24 KD LATE ENDOCYTIC INTERMEDIATE COMPONENT).//3.50E-63//404aa//33%//P32802
C-NT2RP2002942//Homo sapiens mRNA for KIAA0806 protein, complete cds.//0//2090bp//99%//AB018349
C-NT2RP2002974//HOMEOBOX PROTEIN SIX5 (DM LOCUS-ASSOCIATED HOMEODOMAIN PROTEIN HOMOLOG) (FRAGMENT).//8.20E-241//555aa//84%//P70178
C-NT2RP2002976
C-NT2RP2003042//PHOSPHATIDYLCHOLINE-STEROL ACYLTRANSFERASE PRECURSOR (EC 2.3.1.43) (LECITHIN-CHOLESTEROL ACYLTRANSFERASE) (PHOSPHOLIPID-CHOLESTEROL ACYLTRANSFERASE) (FRAGMENT).//2.10E-109//385aa//52%//P53760
C-NT2RP2003179//PUTATIVE SERINE/THREONINE-PROTEIN KINASE EMK (EC 2.7.1.-).//2.60E-67//256aa//49%//Q05512
C-NT2RP2003210//Homo sapiens fatty acid transport protein (FATP) mRNA, complete cds.//9.80E-272//1265bp//98%//AF055899
C-NT2RP2003369//RAS SUPPRESSOR PROTEIN 1 (RSU-1) (RSP-1 PROTEIN) (RSP-1).//5.90E-20//204aa//34%//Q15404
C-NT2RP2003383//Homo sapiens mRNA for KIAA0458 protein, complete cds.//0//2565bp//99%//AB007927
C-NT2RP2003469//GALACTOSE-PROTON SYMPORT (GALACTOSE TRANSPORTER).//1.10E-45//324aa//29%//P37021
C-NT2RP2003545//Homo sapiens STE20-like kinase 3 (mst-3) mRNA, complete cds.//5.40E-48//578bp//71%//AF024636
C-NT2RP2003593
C-NT2RP2003599
C-NT2RP2003655//HYPOTHETICAL 26.3 KD PROTEIN IN OYE2-GND1 INTERGENIC//4.80E-15//93aa//47%//P38869
C-NT2RP2003931
C-NT2RP2004141
C-NT2RP2004179
C-NT2RP2004205//BUTYROPHILIN PRECURSOR (BT).//1.60E-21//276aa//32%//Q62556
C-NT2RP2004447
C-NT2RP2004495//Human transporter protein (g17) mRNA, complete cds.//9.80E-64//642bp//64%//U49082
C-NT2RP2004524
C-NT2RP2004556
C-NT2RP2004606//Human fibroblast collagenase inhibitor mRNA, complete cds.//2.10E-166//768bp//99%//M12670
C-NT2RP2004648//Mouse beta-galactosidase (BGAL) gene, complete cds.//1.20E-33//1136bp//59%//M57734
C-NT2RP2004670//Rattus norvegicus vesicla-associate calmodulin-binding protein mRNA, complete cds.//0//1250bp//86%//L22557
C-NT2RP2004794//HYPOTHETICAL 24.5 KD PROTEIN IN SAP185-BCK1 INTERGENIC REGION.//2.70E-09//203aa//26%//P40857
C-NT2RP2004837
C-NT2RP2004847//ZINC FINGER PROTEIN 135.//8.00E-35//193aa//40%//P52742
C-nnnnnnnnnnnn//Homo sapiens SCG10-like-protein (SCLIP) mRNA, complete cds.//2.90E-170//813bp//98%//AF069709
C-NT2RP2005027
C-NT2RP2005163
C-NT2RP2005181//Mus musculus cationic amino acid transporter (CAT3) mRNA, complete cds.//5.30E-315//2126bp//81%//U70859
C-NT2RP2005247//ZINC-FINGER PROTEIN RFP (RET FINGER PROTEIN).//5.00E-53//296aa//37%//Q62158
C-NT2RP2005425//Homo sapiens mRNA for AKAP450 protein.//0//4341bp//99%//AJ131693
C-NT2RP2005463//PROTEIN PTM1 PRECURSOR.//7.40E-15//284aa//28%//P32857
C-NT2RP2005514
C-NT2RP2005541//N-ACETYLGLUCOSAMINE-6-SULFATASE PRECURSOR (EC 3.1.6.14) (G6S) (GLUCOSAMINE-6-SULFATASE).//4.70E-24//78aa//51%//P15586
C-NT2RP2005632
C-NT2RP2005878//PUTATIVE STEROID DEHYDROGENASE KIK-I (EC 1.1.1.-).//3.60E-55//238aa//50%//O57314
C-NT2RP2005883//DOPAMINE-BETA-MONOOXYGENASE PRECURSOR (EC 1.14.17.1) (DOPAMINE BETA-

NASE I).//1.00E-77//359aa//44%//Q14012
C-Y79AA1001013
C-Y79AA1001056//Homo sapiens MAID protein mRNA, complete cds.//0//1475bp//99%//AF113535
C-Y79AA1001062//TUMOR NECROSIS FACTOR, ALPHA-INDUCED PROTEIN 1, ENDOTHELIAL (B12 PROTEIN).//8.90E-12//132aa//38%//Q13829
C-Y79AA1001090//NUCLEAR FACTOR NF-KAPPA-B P105 SUBUNIT (DNA-BINDING FACTOR KBF1) (EBP- 1) (NF-KAPPA-B1 P84 /NF-KAPPA-B1 P98) [CONTAINS: NUCLEAR FACTOR NF- KAPPA-B P50 SUBUNIT] (FRAGMENT).//4.50E-09//144aa//31%//Q63369
C-Y79AA1001212//Homo sapiens SL15 protein mRNA, complete cds.//6.30E-306//1388bp//99%//AF038961
C-Y79AA1001264//HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II PRECURSOR.//5.10E-106//351aa//58%//Q10005
C-Y79AA1001272//Homo sapiens retinoic acid repressible protein (RARG-1) mRNA, complete cds.//1.50E-183//867bp//98%//AF172066
C-Y79AA1001328//Mus musculus mRNA for Dll3 protein, complete cds.//1.90E-263//1988bp//79%//AB013440
C-Y79AA1001426//ANION EXCHANGE PROTEIN 3 (CARDIAC/BRAIN BAND 3-LIKE PROTEIN) (CAE3/BAE3).//6.20E-66//609aa//31%//P48751
C-Y79AA1001427//Homo sapiens cytochrome b5 reductase 1 (B5R.1) mRNA, complete cds.//0//1588bp//99%//AF169481
C-Y79AA1001430//Homo sapiens mRNA for KIAA0469 protein, complete cds.//0//2943bp//99%//AB007938
C-Y79AA1001523//Homo sapiens transcriptional intermediary factor 1 alpha mRNA, complete cds.//0//2263bp//99%//AF119042
C-Y79AA1001530//Human beta-tubulin gene (5-beta) with ten Alu family members.//0//1920bp//98%//X00734
C-Y79AA1001592
C-Y79AA1001727//CELL SURFACE A33 ANTIGEN PRECURSOR.//1.10E-13//286aa//27%//Q99795
C-Y79AA1001787//PROBABLE CALCIUM-TRANSPORTING ATPASE 9 (EC 3.6.1.38).//1.70E-133//544aa//37%//Q12697
C-Y79AA1001793//Mus musculus mRNA for GSG1, complete cds.//3.70E-126//532bp//78%//D87325
C-Y79AA1001795//Homo sapiens mRNA for GalT4 protein.//2.30E-250//1137bp//99%//Y15061
C-Y79AA1001799//MITOCHONDRIAL RNA SPLICING PROTEIN MSR4.//3.40E-54//182aa//39%//P23500
C-Y79AA1001803//Homo sapiens secretogranin III mRNA, complete cds.//0//1871bp//99%//AF078851
C-Y79AA1001863
C-Y79AA1002022//POLIOVIRUS RECEPTOR HOMOLOG PRECURSOR.//2.20E-06//140aa//26%//P32507
C-Y79AA1002058//Mus musculus Gng3lg mRNA, complete cds.//4.10E-167//1145bp//83%//AF069954
C-Y79AA1002121//HISTONE H1.//4.90E-12//114aa//35%//P35060
C-Y79AA1002129
C-Y79AA1002213//HYPOTHETICAL 52.7 KD PROTEIN C38C10.2 IN CHROMOSOME III.//1.20E-98//262aa//41%//Q03567
C-Y79AA1002334//GLUCOSE REPRESSION MEDIATOR PROTEIN.//1.70E-10//333aa//23%//P14922
C-Y79AA1002373//Mus musculus mRNA for GSG1, complete cds.//7.20E-147//680bp//79%//D87325
C-Y79AA1002376//Rattus norvegicus cytoplasmic dynein intermediate chain 2B mRNA, complete cds.//1.50E-304//1667bp//90%//U39045
C-Y79AA1002378//Homo sapiens zinc finger protein NY-REN-21 antigen mRNA, partial cds.//0//963bp//99%//AF155100
C-Y79AA1002381//Homo sapiens cell cycle related kinase mRNA, complete cds.//0//1791bp//98%//AF035013

Claims

1. Use of an oligonucleotide as a primer for synthesizing the polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 1-829 and 2545, or the complementary strand thereof, wherein said oligonucleotide is complementary to said polynucleotide or the complementary strand thereof and comprises at least 15 nucleotides.

2. A primer set for synthesizing polynucleotides, the primer set comprising an oligo-dT primer and an oligonucleotide complementary to the complementary strand of the polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 1-829 and 2545, wherein said oligonucleotide comprises at least 15 nucleotides.

3. A primer set for synthesizing polynucleotides, the primer set comprising a combination of an oligonucleotide com-

prising a nucleotide sequence complementary to the complementary strand of the polynucleotide comprising a 5'-end nucleotide sequence and an oligonucleotide comprising a nucleotide sequence complementary to the polynucleotide comprising a 3'-end nucleotide sequence, wherein said oligonucleotides comprise at least 15 nucleotides and wherein said combination of 5'-end nucleotide sequence.3'-end nucleotide sequence is selected from the group consisting of:

SEQ ID NO:4 and SEQ ID NO:830
SEQ ID NO:5 and SEQ ID NO:831
SEQ ID NO:6 and SEQ ID NO:832
SEQ ID NO:7 and SEQ ID NO:833
SEQ ID NO:8 and SEQ ID NO:834
SEQ ID NO:9 and SEQ ID NO:835
SEQ ID NO:11 and SEQ ID NO:836
SEQ ID NO:12 and SEQ ID NO:837
SEQ ID NO:13 and SEQ ID NO:838
SEQ ID NO:14 and SEQ ID NO:839
SEQ ID NO:15 and SEQ ID NO:840
SEQ ID NO:16 and SEQ ID NO:841
SEQ ID NO:17 and SEQ ID NO:842
SEQ ID NO:18 and SEQ ID NO:843
SEQ ID NO:20 and SEQ ID NO:844
SEQ ID NO:22 and SEQ ID NO:845
SEQ ID NO:23 and SEQ ID NO:846
SEQ ID NO:24 and SEQ ID NO:847
SEQ ID NO:25 and SEQ ID NO:848
SEQ ID NO:26 and SEQ ID NO:849
SEQ ID NO:27 and SEQ ID NO:850
SEQ ID NO:28 and SEQ ID NO:851
SEQ ID NO:29 and SEQ ID NO:852
SEQ ID NO:30 and SEQ ID NO:853
SEQ ID NO:31 and SEQ ID NO:854
SEQ ID NO:32 and SEQ ID NO:855
SEQ ID NO:33 and SEQ ID NO:856
SEQ ID NO:34 and SEQ ID NO:857
SEQ ID NO:35 and SEQ ID NO:858
SEQ ID NO:36 and SEQ ID NO:859
SEQ ID NO:37 and SEQ ID NO:860
SEQ ID NO:39 and SEQ ID NO:861
SEQ ID NO:40 and SEQ ID NO:862
SEQ ID NO:41 and SEQ ID NO:863
SEQ ID NO:42 and SEQ ID NO:864
SEQ ID NO:44 and SEQ ID NO:865
SEQ ID NO:45 and SEQ ID NO:866
SEQ ID NO:46 and SEQ ID NO:867
SEQ ID NO:47 and SEQ ID NO:868

SEQ ID NO:48 and SEQ ID NO:869
SEQ ID NO:49 and SEQ ID NO:870
SEQ ID NO:50 and SEQ ID NO:871
SEQ ID NO:52 and SEQ ID NO:872
SEQ ID NO:53 and SEQ ID NO:873
SEQ ID NO:54 and SEQ ID NO:874
SEQ ID NO:55 and SEQ ID NO:875
SEQ ID NO:56 and SEQ ID NO:876
SEQ ID NO:57 and SEQ ID NO:877
SEQ ID NO:58 and SEQ ID NO:878
SEQ ID NO:59 and SEQ ID NO:879
SEQ ID NO:60 and SEQ ID NO:880
SEQ ID NO:61 and SEQ ID NO:881
SEQ ID NO:62 and SEQ ID NO:882
SEQ ID NO:63 and SEQ ID NO:883
SEQ ID NO:64 and SEQ ID NO:884
SEQ ID NO:65 and SEQ ID NO:885
SEQ ID NO:66 and SEQ ID NO:886
SEQ ID NO:67 and SEQ ID NO:887
SEQ ID NO:69 and SEQ ID NO:888
SEQ ID NO:70 and SEQ ID NO:889
SEQ ID NO:71 and SEQ ID NO:890
SEQ ID NO:72 and SEQ ID NO:891
SEQ ID NO:73 and SEQ ID NO:892
SEQ ID NO:74 and SEQ ID NO:893
SEQ ID NO:75 and SEQ ID NO:894
SEQ ID NO:76 and SEQ ID NO:895
SEQ ID NO:77 and SEQ ID NO:896
SEQ ID NO:78 and SEQ ID NO:897
SEQ ID NO:79 and SEQ ID NO:898
SEQ ID NO:80 and SEQ ID NO:899
SEQ ID NO:81 and SEQ ID NO:900
SEQ ID NO:82 and SEQ ID NO:901
SEQ ID NO:83 and SEQ ID NO:902
SEQ ID NO:84 and SEQ ID NO:903
SEQ ID NO:85 and SEQ ID NO:904
SEQ ID NO:86 and SEQ ID NO:905
SEQ ID NO:87 and SEQ ID NO:906
SEQ ID NO:88 and SEQ ID NO:907
SEQ ID NO:89 and SEQ ID NO:908
SEQ ID NO:90 and SEQ ID NO:909
SEQ ID NO:91 and SEQ ID NO:910
SEQ ID NO:92 and SEQ ID NO:911
SEQ ID NO:93 and SEQ ID NO:912
SEQ ID NO:94 and SEQ ID NO:913
SEQ ID NO:95 and SEQ ID NO:914
SEQ ID NO:96 and SEQ ID NO:915
SEQ ID NO:97 and SEQ ID NO:916
SEQ ID NO:98 and SEQ ID NO:917
SEQ ID NO:99 and SEQ ID NO:918
SEQ ID NO:100 and SEQ ID NO:919
SEQ ID NO:101 and SEQ ID NO:920

SEQ ID NO:102 and SEQ ID NO:921
SEQ ID NO:103 and SEQ ID NO:922
SEQ ID NO:104 and SEQ ID NO:923
SEQ ID NO:105 and SEQ ID NO:924
SEQ ID NO:106 and SEQ ID NO:925
SEQ ID NO:107 and SEQ ID NO:926
SEQ ID NO:108 and SEQ ID NO:927
SEQ ID NO:109 and SEQ ID NO:928
SEQ ID NO:110 and SEQ ID NO:929
SEQ ID NO:111 and SEQ ID NO:930
SEQ ID NO:112 and SEQ ID NO:931
SEQ ID NO:113 and SEQ ID NO:932
SEQ ID NO:114 and SEQ ID NO:933
SEQ ID NO:115 and SEQ ID NO:934
SEQ ID NO:116 and SEQ ID NO:935
SEQ ID NO:117 and SEQ ID NO:936
SEQ ID NO:118 and SEQ ID NO:937
SEQ ID NO:119 and SEQ ID NO:938
SEQ ID NO:120 and SEQ ID NO:939
SEQ ID NO:122 and SEQ ID NO:940
SEQ ID NO:123 and SEQ ID NO:941
SEQ ID NO:124 and SEQ ID NO:942
SEQ ID NO:125 and SEQ ID NO:943
SEQ ID NO:126 and SEQ ID NO:944
SEQ ID NO:127 and SEQ ID NO:945
SEQ ID NO:128 and SEQ ID NO:946
SEQ ID NO:129 and SEQ ID NO:947
SEQ ID NO:130 and SEQ ID NO:948
SEQ ID NO:131 and SEQ ID NO:949
SEQ ID NO:132 and SEQ ID NO:950
SEQ ID NO:133 and SEQ ID NO:951
SEQ ID NO:134 and SEQ ID NO:952
SEQ ID NO:135 and SEQ ID NO:953
SEQ ID NO:136 and SEQ ID NO:954
SEQ ID NO:137 and SEQ ID NO:955
SEQ ID NO:138 and SEQ ID NO:956
SEQ ID NO:139 and SEQ ID NO:957
SEQ ID NO:140 and SEQ ID NO:958
SEQ ID NO:141 and SEQ ID NO:959
SEQ ID NO:142 and SEQ ID NO:960
SEQ ID NO:143 and SEQ ID NO:961
SEQ ID NO:144 and SEQ ID NO:962
SEQ ID NO:145 and SEQ ID NO:963
SEQ ID NO:146 and SEQ ID NO:964
SEQ ID NO:147 and SEQ ID NO:965
SEQ ID NO:148 and SEQ ID NO:966
SEQ ID NO:149 and SEQ ID NO:967
SEQ ID NO:150 and SEQ ID NO:968
SEQ ID NO:151 and SEQ ID NO:969
SEQ ID NO:152 and SEQ ID NO:970
SEQ ID NO:153 and SEQ ID NO:971
SEQ ID NO:154 and SEQ ID NO:972

SEQ ID NO:155 and SEQ ID NO:973
SEQ ID NO:156 and SEQ ID NO:974
SEQ ID NO:157 and SEQ ID NO:975
SEQ ID NO:158 and SEQ ID NO:976
SEQ ID NO:159 and SEQ ID NO:977
SEQ ID NO:160 and SEQ ID NO:978
SEQ ID NO:161 and SEQ ID NO:979
SEQ ID NO:162 and SEQ ID NO:980
SEQ ID NO:163 and SEQ ID NO:981
SEQ ID NO:164 and SEQ ID NO:982
SEQ ID NO:165 and SEQ ID NO:983
SEQ ID NO:166 and SEQ ID NO:984
SEQ ID NO:167 and SEQ ID NO:985
SEQ ID NO:168 and SEQ ID NO:986
SEQ ID NO:169 and SEQ ID NO:987
SEQ ID NO:170 and SEQ ID NO:988
SEQ ID NO:171 and SEQ ID NO:989
SEQ ID NO:172 and SEQ ID NO:990
SEQ ID NO:173 and SEQ ID NO:991
SEQ ID NO:174 and SEQ ID NO:992
SEQ ID NO:175 and SEQ ID NO:993
SEQ ID NO:176 and SEQ ID NO:994
SEQ ID NO:177 and SEQ ID NO:995
SEQ ID NO:178 and SEQ ID NO:996
SEQ ID NO:179 and SEQ ID NO:997
SEQ ID NO:180 and SEQ ID NO:998
SEQ ID NO:181 and SEQ ID NO:999
SEQ ID NO:182 and SEQ ID NO:1000
SEQ ID NO:183 and SEQ ID NO:1001
SEQ ID NO:184 and SEQ ID NO:1002
SEQ ID NO:185 and SEQ ID NO:1003
SEQ ID NO:186 and SEQ ID NO:1004
SEQ ID NO:187 and SEQ ID NO:1005
SEQ ID NO:188 and SEQ ID NO:1006
SEQ ID NO:189 and SEQ ID NO:1007
SEQ ID NO:190 and SEQ ID NO:1008
SEQ ID NO:191 and SEQ ID NO:1009
SEQ ID NO:193 and SEQ ID NO:1010
SEQ ID NO:194 and SEQ ID NO:1011
SEQ ID NO:195 and SEQ ID NO:1012
SEQ ID NO:196 and SEQ ID NO:1013
SEQ ID NO:197 and SEQ ID NO:1014
SEQ ID NO:198 and SEQ ID NO:1015
SEQ ID NO:199 and SEQ ID NO:1016
SEQ ID NO:200 and SEQ ID NO:1017
SEQ ID NO:201 and SEQ ID NO:1018
SEQ ID NO:202 and SEQ ID NO:1019
SEQ ID NO:203 and SEQ ID NO:1020
SEQ ID NO:204 and SEQ ID NO:1021
SEQ ID NO:205 and SEQ ID NO:1022
SEQ ID NO:207 and SEQ ID NO:1023
SEQ ID NO:208 and SEQ ID NO:1024

SEQ ID NO:209 and SEQ ID NO:1025
SEQ ID NO:210 and SEQ ID NO:1026
SEQ ID NO:211 and SEQ ID NO:1027
SEQ ID NO:212 and SEQ ID NO:1028
SEQ ID NO:213 and SEQ ID NO:1029
SEQ ID NO:214 and SEQ ID NO:1030
SEQ ID NO:215 and SEQ ID NO:1031
SEQ ID NO:216 and SEQ ID NO:1032
SEQ ID NO:248 and SEQ ID NO:1033
SEQ ID NO:249 and SEQ ID NO:1034
SEQ ID NO:250 and SEQ ID NO:1035
SEQ ID NO:251 and SEQ ID NO:1036
SEQ ID NO:252 and SEQ ID NO:1037
SEQ ID NO:253 and SEQ ID NO:1038
SEQ ID NO:254 and SEQ ID NO:1039
SEQ ID NO:255 and SEQ ID NO:1040
SEQ ID NO:256 and SEQ ID NO:1041
SEQ ID NO:257 and SEQ ID NO:1042
SEQ ID NO:258 and SEQ ID NO:1043
SEQ ID NO:259 and SEQ ID NO:1044
SEQ ID NO:260 and SEQ ID NO:1045
SEQ ID NO:261 and SEQ ID NO:1046
SEQ ID NO:262 and SEQ ID NO:1047
SEQ ID NO:263 and SEQ ID NO:1048
SEQ ID NO:264 and SEQ ID NO:1049
SEQ ID NO:265 and SEQ ID NO:1050
SEQ ID NO:266 and SEQ ID NO:1051
SEQ ID NO:267 and SEQ ID NO:1052
SEQ ID NO:268 and SEQ ID NO:1053
SEQ ID NO:291 and SEQ ID NO:1054
SEQ ID NO:292 and SEQ ID NO:1055
SEQ ID NO:293 and SEQ ID NO:1056
SEQ ID NO:294 and SEQ ID NO:1057
SEQ ID NO:295 and SEQ ID NO:1058
SEQ ID NO:296 and SEQ ID NO:1059
SEQ ID NO:297 and SEQ ID NO:1060
SEQ ID NO:298 and SEQ ID NO:1061
SEQ ID NO:300 and SEQ ID NO:1062
SEQ ID NO:301 and SEQ ID NO:1063
SEQ ID NO:302 and SEQ ID NO:1064
SEQ ID NO:304 and SEQ ID NO:1065
SEQ ID NO:305 and SEQ ID NO:1066
SEQ ID NO:306 and SEQ ID NO:1067
SEQ ID NO:307 and SEQ ID NO:1068
SEQ ID NO:308 and SEQ ID NO:1069
SEQ ID NO:309 and SEQ ID NO:1070
SEQ ID NO:310 and SEQ ID NO:1071
SEQ ID NO:311 and SEQ ID NO:1072
SEQ ID NO:312 and SEQ ID NO:1073
SEQ ID NO:313 and SEQ ID NO:1074
SEQ ID NO:314 and SEQ ID NO:1075
SEQ ID NO:315 and SEQ ID NO:1076

SEQ ID NO:317 and SEQ ID NO:1077
SEQ ID NO:318 and SEQ ID NO:1078
SEQ ID NO:319 and SEQ ID NO:1079
SEQ ID NO:320 and SEQ ID NO:1080
SEQ ID NO:321 and SEQ ID NO:1081
SEQ ID NO:322 and SEQ ID NO:1082
SEQ ID NO:323 and SEQ ID NO:1083
SEQ ID NO:324 and SEQ ID NO:1084
SEQ ID NO:325 and SEQ ID NO:1085
SEQ ID NO:326 and SEQ ID NO:1086
SEQ ID NO:327 and SEQ ID NO:1087
SEQ ID NO:328 and SEQ ID NO:1088
SEQ ID NO:329 and SEQ ID NO:1089
SEQ ID NO:330 and SEQ ID NO:1090
SEQ ID NO:331 and SEQ ID NO:1091
SEQ ID NO:332 and SEQ ID NO:1092
SEQ ID NO:333 and SEQ ID NO:1093
SEQ ID NO:334 and SEQ ID NO:1094
SEQ ID NO:335 and SEQ ID NO:1095
SEQ ID NO:336 and SEQ ID NO:1096
SEQ ID NO:337 and SEQ ID NO:1097
SEQ ID NO:338 and SEQ ID NO:1098
SEQ ID NO:339 and SEQ ID NO:1099
SEQ ID NO:340 and SEQ ID NO:1100
SEQ ID NO:341 and SEQ ID NO:1101
SEQ ID NO:342 and SEQ ID NO:1102
SEQ ID NO:343 and SEQ ID NO:1103
SEQ ID NO:344 and SEQ ID NO:1104
SEQ ID NO:345 and SEQ ID NO:1105
SEQ ID NO:347 and SEQ ID NO:1106
SEQ ID NO:348 and SEQ ID NO:1107
SEQ ID NO:349 and SEQ ID NO:1108
SEQ ID NO:350 and SEQ ID NO:1109
SEQ ID NO:351 and SEQ ID NO:1110
SEQ ID NO:352 and SEQ ID NO:1111
SEQ ID NO:353 and SEQ ID NO:1112
SEQ ID NO:354 and SEQ ID NO:1113
SEQ ID NO:355 and SEQ ID NO:1114
SEQ ID NO:356 and SEQ ID NO:1115
SEQ ID NO:357 and SEQ ID NO:1116
SEQ ID NO:358 and SEQ ID NO:1117
SEQ ID NO:359 and SEQ ID NO:1118
SEQ ID NO:360 and SEQ ID NO:1119
SEQ ID NO:361 and SEQ ID NO:1120
SEQ ID NO:362 and SEQ ID NO:1121
SEQ ID NO:363 and SEQ ID NO:1122
SEQ ID NO:364 and SEQ ID NO:1123
SEQ ID NO:365 and SEQ ID NO:1124
SEQ ID NO:366 and SEQ ID NO:1125
SEQ ID NO:367 and SEQ ID NO:1126
SEQ ID NO:368 and SEQ ID NO:1127
SEQ ID NO:369 and SEQ ID NO:1128

SEQ ID NO:370 and SEQ ID NO:1129
SEQ ID NO:371 and SEQ ID NO:1130
SEQ ID NO:372 and SEQ ID NO:1131
SEQ ID NO:373 and SEQ ID NO:1132
SEQ ID NO:374 and SEQ ID NO:1133
SEQ ID NO:375 and SEQ ID NO:1134
SEQ ID NO:376 and SEQ ID NO:1135
SEQ ID NO:377 and SEQ ID NO:1136
SEQ ID NO:378 and SEQ ID NO:1137
SEQ ID NO:379 and SEQ ID NO:1138
SEQ ID NO:380 and SEQ ID NO:1139
SEQ ID NO:381 and SEQ ID NO:1140
SEQ ID NO:382 and SEQ ID NO:1141
SEQ ID NO:383 and SEQ ID NO:1142
SEQ ID NO:384 and SEQ ID NO:1143
SEQ ID NO:385 and SEQ ID NO:1144
SEQ ID NO:386 and SEQ ID NO:1145
SEQ ID NO:387 and SEQ ID NO:1146
SEQ ID NO:388 and SEQ ID NO:1147
SEQ ID NO:389 and SEQ ID NO:1148
SEQ ID NO:390 and SEQ ID NO:1149
SEQ ID NO:391 and SEQ ID NO:1150
SEQ ID NO:392 and SEQ ID NO:1151
SEQ ID NO:393 and SEQ ID NO:1152
SEQ ID NO:394 and SEQ ID NO:1153
SEQ ID NO:395 and SEQ ID NO:1154
SEQ ID NO:396 and SEQ ID NO:1155
SEQ ID NO:397 and SEQ ID NO:1156
SEQ ID NO:398 and SEQ ID NO:1157
SEQ ID NO:399 and SEQ ID NO:1158
SEQ ID NO:400 and SEQ ID NO:1159
SEQ ID NO:401 and SEQ ID NO:1160
SEQ ID NO:402 and SEQ ID NO:1161
SEQ ID NO:403 and SEQ ID NO:1162
SEQ ID NO:404 and SEQ ID NO:1163
SEQ ID NO:405 and SEQ ID NO:1164
SEQ ID NO:406 and SEQ ID NO:1165
SEQ ID NO:407 and SEQ ID NO:1166
SEQ ID NO:408 and SEQ ID NO:1167
SEQ ID NO:409 and SEQ ID NO:1168
SEQ ID NO:410 and SEQ ID NO:1169
SEQ ID NO:411 and SEQ ID NO:1170
SEQ ID NO:412 and SEQ ID NO:1171
SEQ ID NO:413 and SEQ ID NO:1172
SEQ ID NO:414 and SEQ ID NO:1173
SEQ ID NO:415 and SEQ ID NO:1174
SEQ ID NO:416 and SEQ ID NO:1175
SEQ ID NO:417 and SEQ ID NO:1176
SEQ ID NO:419 and SEQ ID NO:1177
SEQ ID NO:420 and SEQ ID NO:1178
SEQ ID NO:421 and SEQ ID NO:1179
SEQ ID NO:422 and SEQ ID NO:1180

SEQ ID NO:423 and SEQ ID NO:1181
SEQ ID NO:424 and SEQ ID NO:1182
SEQ ID NO:425 and SEQ ID NO:1183
SEQ ID NO:426 and SEQ ID NO:1184
SEQ ID NO:427 and SEQ ID NO:1185
SEQ ID NO:428 and SEQ ID NO:1186
SEQ ID NO:429 and SEQ ID NO:1187
SEQ ID NO:430 and SEQ ID NO:1188
SEQ ID NO:431 and SEQ ID NO:1189
SEQ ID NO:432 and SEQ ID NO:1190
SEQ ID NO:433 and SEQ ID NO:1191
SEQ ID NO:434 and SEQ ID NO:1192
SEQ ID NO:435 and SEQ ID NO:1193
SEQ ID NO:436 and SEQ ID NO:1194
SEQ ID NO:437 and SEQ ID NO:1195
SEQ ID NO:438 and SEQ ID NO:1196
SEQ ID NO:439 and SEQ ID NO:1197
SEQ ID NO:440 and SEQ ID NO:1198
SEQ ID NO:441 and SEQ ID NO:1199
SEQ ID NO:442 and SEQ ID NO:1200
SEQ ID NO:443 and SEQ ID NO:1201
SEQ ID NO:444 and SEQ ID NO:1202
SEQ ID NO:446 and SEQ ID NO:1203
SEQ ID NO:447 and SEQ ID NO:1204
SEQ ID NO:448 and SEQ ID NO:1205
SEQ ID NO:450 and SEQ ID NO:1206
SEQ ID NO:451 and SEQ ID NO:1207
SEQ ID NO:452 and SEQ ID NO:1208
SEQ ID NO:453 and SEQ ID NO:1209
SEQ ID NO:454 and SEQ ID NO:1210
SEQ ID NO:455 and SEQ ID NO:1211
SEQ ID NO:456 and SEQ ID NO:1212
SEQ ID NO:457 and SEQ ID NO:1213
SEQ ID NO:458 and SEQ ID NO:1214
SEQ ID NO:459 and SEQ ID NO:1215
SEQ ID NO:460 and SEQ ID NO:1216
SEQ ID NO:461 and SEQ ID NO:1217
SEQ ID NO:462 and SEQ ID NO:1218
SEQ ID NO:463 and SEQ ID NO:1219
SEQ ID NO:464 and SEQ ID NO:1220
SEQ ID NO:465 and SEQ ID NO:1221
SEQ ID NO:466 and SEQ ID NO:1222
SEQ ID NO:467 and SEQ ID NO:1223
SEQ ID NO:468 and SEQ ID NO:1224
SEQ ID NO:469 and SEQ ID NO:1225
SEQ ID NO:470 and SEQ ID NO:1226
SEQ ID NO:471 and SEQ ID NO:1227
SEQ ID NO:472 and SEQ ID NO:1228
SEQ ID NO:473 and SEQ ID NO:1229
SEQ ID NO:474 and SEQ ID NO:1230
SEQ ID NO:475 and SEQ ID NO:1231
SEQ ID NO:476 and SEQ ID NO:1232

SEQ ID NO:477 and SEQ ID NO:1233
SEQ ID NO:478 and SEQ ID NO:1234
SEQ ID NO:479 and SEQ ID NO:1235
SEQ ID NO:480 and SEQ ID NO:1236
SEQ ID NO:481 and SEQ ID NO:1237
SEQ ID NO:482 and SEQ ID NO:1238
SEQ ID NO:483 and SEQ ID NO:1239
SEQ ID NO:484 and SEQ ID NO:1240
SEQ ID NO:485 and SEQ ID NO:1241
SEQ ID NO:486 and SEQ ID NO:1242
SEQ ID NO:487 and SEQ ID NO:1243
SEQ ID NO:488 and SEQ ID NO:1244
SEQ ID NO:489 and SEQ ID NO:1245
SEQ ID NO:490 and SEQ ID NO:1246
SEQ ID NO:492 and SEQ ID NO:1247
SEQ ID NO:493 and SEQ ID NO:1248
SEQ ID NO:494 and SEQ ID NO:1249
SEQ ID NO:496 and SEQ ID NO:1250
SEQ ID NO:497 and SEQ ID NO:1251
SEQ ID NO:498 and SEQ ID NO:1252
SEQ ID NO:499 and SEQ ID NO:1253
SEQ ID NO:500 and SEQ ID NO:1254
SEQ ID NO:501 and SEQ ID NO:1255
SEQ ID NO:502 and SEQ ID NO:1256
SEQ ID NO:503 and SEQ ID NO:1257
SEQ ID NO:504 and SEQ ID NO:1258
SEQ ID NO:505 and SEQ ID NO:1259
SEQ ID NO:506 and SEQ ID NO:1260
SEQ ID NO:507 and SEQ ID NO:1261
SEQ ID NO:508 and SEQ ID NO:1262
SEQ ID NO:509 and SEQ ID NO:1263
SEQ ID NO:510 and SEQ ID NO:1264
SEQ ID NO:511 and SEQ ID NO:1265
SEQ ID NO:513 and SEQ ID NO:1266
SEQ ID NO:514 and SEQ ID NO:1267
SEQ ID NO:515 and SEQ ID NO:1268
SEQ ID NO:516 and SEQ ID NO:1269
SEQ ID NO:517 and SEQ ID NO:1270
SEQ ID NO:518 and SEQ ID NO:1271
SEQ ID NO:519 and SEQ ID NO:1272
SEQ ID NO:520 and SEQ ID NO:1273
SEQ ID NO:521 and SEQ ID NO:1274
SEQ ID NO:522 and SEQ ID NO:1275
SEQ ID NO:523 and SEQ ID NO:1276
SEQ ID NO:524 and SEQ ID NO:1277
SEQ ID NO:525 and SEQ ID NO:1278
SEQ ID NO:526 and SEQ ID NO:1279
SEQ ID NO:527 and SEQ ID NO:1280
SEQ ID NO:528 and SEQ ID NO:1281
SEQ ID NO:529 and SEQ ID NO:1282
SEQ ID NO:530 and SEQ ID NO:1283
SEQ ID NO:531 and SEQ ID NO:1284

SEQ ID NO:532 and SEQ ID NO:1285
SEQ ID NO:533 and SEQ ID NO:1286
SEQ ID NO:534 and SEQ ID NO:1287
SEQ ID NO:535 and SEQ ID NO:1288
SEQ ID NO:536 and SEQ ID NO:1289
SEQ ID NO:537 and SEQ ID NO:1290
SEQ ID NO:538 and SEQ ID NO:1291
SEQ ID NO:539 and SEQ ID NO:1292
SEQ ID NO:540 and SEQ ID NO:1293
SEQ ID NO:541 and SEQ ID NO:1294
SEQ ID NO:542 and SEQ ID NO:1295
SEQ ID NO:543 and SEQ ID NO:1296
SEQ ID NO:544 and SEQ ID NO:1297
SEQ ID NO:545 and SEQ ID NO:1298
SEQ ID NO:546 and SEQ ID NO:1299
SEQ ID NO:547 and SEQ ID NO:1300
SEQ ID NO:548 and SEQ ID NO:1301
SEQ ID NO:549 and SEQ ID NO:1302
SEQ ID NO:550 and SEQ ID NO:1303
SEQ ID NO:551 and SEQ ID NO:1304
SEQ ID NO:552 and SEQ ID NO:1305
SEQ ID NO:553 and SEQ ID NO:1306
SEQ ID NO:554 and SEQ ID NO:1307
SEQ ID NO:555 and SEQ ID NO:1308
SEQ ID NO:556 and SEQ ID NO:1309
SEQ ID NO:557 and SEQ ID NO:1310
SEQ ID NO:558 and SEQ ID NO:1311
SEQ ID NO:559 and SEQ ID NO:1312
SEQ ID NO:560 and SEQ ID NO:1313
SEQ ID NO:561 and SEQ ID NO:1314
SEQ ID NO:562 and SEQ ID NO:1315
SEQ ID NO:563 and SEQ ID NO:1316
SEQ ID NO:564 and SEQ ID NO:1317
SEQ ID NO:565 and SEQ ID NO:1318
SEQ ID NO:566 and SEQ ID NO:1319
SEQ ID NO:567 and SEQ ID NO:1320
SEQ ID NO:569 and SEQ ID NO:1321
SEQ ID NO:570 and SEQ ID NO:1322
SEQ ID NO:571 and SEQ ID NO:1323
SEQ ID NO:572 and SEQ ID NO:1324
SEQ ID NO:573 and SEQ ID NO:1325
SEQ ID NO:574 and SEQ ID NO:1326
SEQ ID NO:575 and SEQ ID NO:1327
SEQ ID NO:576 and SEQ ID NO:1328
SEQ ID NO:577 and SEQ ID NO:1329
SEQ ID NO:579 and SEQ ID NO:1330
SEQ ID NO:580 and SEQ ID NO:1331
SEQ ID NO:581 and SEQ ID NO:1332
SEQ ID NO:582 and SEQ ID NO:1333
SEQ ID NO:583 and SEQ ID NO:1334
SEQ ID NO:584 and SEQ ID NO:1335
SEQ ID NO:585 and SEQ ID NO:1336

SEQ ID NO:586 and SEQ ID NO:1337
SEQ ID NO:587 and SEQ ID NO:1338
SEQ ID NO:588 and SEQ ID NO:1339
SEQ ID NO:589 and SEQ ID NO:1340
SEQ ID NO:590 and SEQ ID NO:1341
SEQ ID NO:591 and SEQ ID NO:1342
SEQ ID NO:592 and SEQ ID NO:1343
SEQ ID NO:593 and SEQ ID NO:1344
SEQ ID NO:594 and SEQ ID NO:1345
SEQ ID NO:595 and SEQ ID NO:1346
SEQ ID NO:596 and SEQ ID NO:1347
SEQ ID NO:597 and SEQ ID NO:1348
SEQ ID NO:598 and SEQ ID NO:1349
SEQ ID NO:599 and SEQ ID NO:1350
SEQ ID NO:600 and SEQ ID NO:1351
SEQ ID NO:601 and SEQ ID NO:1352
SEQ ID NO:602 and SEQ ID NO:1353
SEQ ID NO:603 and SEQ ID NO:1354
SEQ ID NO:604 and SEQ ID NO:1355
SEQ ID NO:605 and SEQ ID NO:1356
SEQ ID NO:606 and SEQ ID NO:1357
SEQ ID NO:607 and SEQ ID NO:1358
SEQ ID NO:608 and SEQ ID NO:1359
SEQ ID NO:609 and SEQ ID NO:1360
SEQ ID NO:610 and SEQ ID NO:1361
SEQ ID NO:611 and SEQ ID NO:1362
SEQ ID NO:612 and SEQ ID NO:1363
SEQ ID NO:613 and SEQ ID NO:1364
SEQ ID NO:614 and SEQ ID NO:1365
SEQ ID NO:615 and SEQ ID NO:1366
SEQ ID NO:616 and SEQ ID NO:1367
SEQ ID NO:617 and SEQ ID NO:1368
SEQ ID NO:618 and SEQ ID NO:1369
SEQ ID NO:619 and SEQ ID NO:1370
SEQ ID NO:620 and SEQ ID NO:1371
SEQ ID NO:621 and SEQ ID NO:1372
SEQ ID NO:622 and SEQ ID NO:1373
SEQ ID NO:623 and SEQ ID NO:1374
SEQ ID NO:624 and SEQ ID NO:1375
SEQ ID NO:625 and SEQ ID NO:1376
SEQ ID NO:626 and SEQ ID NO:1377
SEQ ID NO:627 and SEQ ID NO:1378
SEQ ID NO:628 and SEQ ID NO:1379
SEQ ID NO:629 and SEQ ID NO:1380
SEQ ID NO:630 and SEQ ID NO:1381
SEQ ID NO:631 and SEQ ID NO:1382
SEQ ID NO:632 and SEQ ID NO:1383
SEQ ID NO:633 and SEQ ID NO:1384
SEQ ID NO:634 and SEQ ID NO:1385
SEQ ID NO:635 and SEQ ID NO:1386
SEQ ID NO:636 and SEQ ID NO:1387
SEQ ID NO:637 and SEQ ID NO:1388

SEQ ID NO:638 and SEQ ID NO:1389
SEQ ID NO:639 and SEQ ID NO:1390
SEQ ID NO:640 and SEQ ID NO:1391
SEQ ID NO:641 and SEQ ID NO:1392
SEQ ID NO:642 and SEQ ID NO:1393
SEQ ID NO:643 and SEQ ID NO:1394
SEQ ID NO:644 and SEQ ID NO:1395
SEQ ID NO:645 and SEQ ID NO:1396
SEQ ID NO:646 and SEQ ID NO:1397
SEQ ID NO:647 and SEQ ID NO:1398
SEQ ID NO:648 and SEQ ID NO:1399
SEQ ID NO:649 and SEQ ID NO:1400
SEQ ID NO:650 and SEQ ID NO:1401
SEQ ID NO:651 and SEQ ID NO:1402
SEQ ID NO:652 and SEQ ID NO:1403
SEQ ID NO:653 and SEQ ID NO:1404
SEQ ID NO:654 and SEQ ID NO:1405
SEQ ID NO:655 and SEQ ID NO:1406
SEQ ID NO:656 and SEQ ID NO:1407
SEQ ID NO:657 and SEQ ID NO:1408
SEQ ID NO:658 and SEQ ID NO:1409
SEQ ID NO:659 and SEQ ID NO:1410
SEQ ID NO:660 and SEQ ID NO:1411
SEQ ID NO:661 and SEQ ID NO:1412
SEQ ID NO:662 and SEQ ID NO:1413
SEQ ID NO:663 and SEQ ID NO:1414
SEQ ID NO:664 and SEQ ID NO:1415
SEQ ID NO:665 and SEQ ID NO:1416
SEQ ID NO:666 and SEQ ID NO:1417
SEQ ID NO:667 and SEQ ID NO:1418
SEQ ID NO:668 and SEQ ID NO:1419
SEQ ID NO:669 and SEQ ID NO:1420
SEQ ID NO:670 and SEQ ID NO:1421
SEQ ID NO:671 and SEQ ID NO:1422
SEQ ID NO:672 and SEQ ID NO:1423
SEQ ID NO:673 and SEQ ID NO:1424
SEQ ID NO:674 and SEQ ID NO:1425
SEQ ID NO:675 and SEQ ID NO:1426
SEQ ID NO:676 and SEQ ID NO:1427
SEQ ID NO:677 and SEQ ID NO:1428
SEQ ID NO:678 and SEQ ID NO:1429
SEQ ID NO:679 and SEQ ID NO:1430
SEQ ID NO:680 and SEQ ID NO:1431
SEQ ID NO:681 and SEQ ID NO:1432
SEQ ID NO:682 and SEQ ID NO:1433
SEQ ID NO:683 and SEQ ID NO:1434
SEQ ID NO:684 and SEQ ID NO:1435
SEQ ID NO:685 and SEQ ID NO:1436
SEQ ID NO:686 and SEQ ID NO:1437
SEQ ID NO:687 and SEQ ID NO:1438
SEQ ID NO:688 and SEQ ID NO:1439
SEQ ID NO:689 and SEQ ID NO:1440

SEQ ID NO:690 and SEQ ID NO:1441
SEQ ID NO:691 and SEQ ID NO:1442
SEQ ID NO:692 and SEQ ID NO:1443
SEQ ID NO:693 and SEQ ID NO:1444
SEQ ID NO:694 and SEQ ID NO:1445
SEQ ID NO:695 and SEQ ID NO:1446
SEQ ID NO:696 and SEQ ID NO:1447
SEQ ID NO:697 and SEQ ID NO:1448
SEQ ID NO:698 and SEQ ID NO:1449
SEQ ID NO:699 and SEQ ID NO:1450
SEQ ID NO:700 and SEQ ID NO:1451
SEQ ID NO:701 and SEQ ID NO:1452
SEQ ID NO:702 and SEQ ID NO:1453
SEQ ID NO:703 and SEQ ID NO:1454
SEQ ID NO:704 and SEQ ID NO:1455
SEQ ID NO:705 and SEQ ID NO:1456
SEQ ID NO:706 and SEQ ID NO:1457
SEQ ID NO:707 and SEQ ID NO:1458
SEQ ID NO:708 and SEQ ID NO:1459
SEQ ID NO:709 and SEQ ID NO:1460
SEQ ID NO:710 and SEQ ID NO:1461
SEQ ID NO:711 and SEQ ID NO:1462
SEQ ID NO:712 and SEQ ID NO:1463
SEQ ID NO:713 and SEQ ID NO:1464
SEQ ID NO:714 and SEQ ID NO:1465
SEQ ID NO:715 and SEQ ID NO:1466
SEQ ID NO:716 and SEQ ID NO:1467
SEQ ID NO:717 and SEQ ID NO:1468
SEQ ID NO:718 and SEQ ID NO:1469
SEQ ID NO:719 and SEQ ID NO:1470
SEQ ID NO:720 and SEQ ID NO:1471
SEQ ID NO:721 and SEQ ID NO:1472
SEQ ID NO:722 and SEQ ID NO:1473
SEQ ID NO:723 and SEQ ID NO:1474
SEQ ID NO:724 and SEQ ID NO:1475
SEQ ID NO:725 and SEQ ID NO:1476
SEQ ID NO:726 and SEQ ID NO:1477
SEQ ID NO:727 and SEQ ID NO:1478
SEQ ID NO:729 and SEQ ID NO:1479
SEQ ID NO:730 and SEQ ID NO:1480
SEQ ID NO:731 and SEQ ID NO:1481
SEQ ID NO:732 and SEQ ID NO:1482
SEQ ID NO:733 and SEQ ID NO:1483
SEQ ID NO:734 and SEQ ID NO:1484
SEQ ID NO:735 and SEQ ID NO:1485
SEQ ID NO:736 and SEQ ID NO:1486
SEQ ID NO:737 and SEQ ID NO:1487
SEQ ID NO:738 and SEQ ID NO:1488
SEQ ID NO:739 and SEQ ID NO:1489
SEQ ID NO:740 and SEQ ID NO:1490
SEQ ID NO:741 and SEQ ID NO:1491
SEQ ID NO:745 and SEQ ID NO:1492

SEQ ID NO:746 and SEQ ID NO:1493
SEQ ID NO:747 and SEQ ID NO:1494
SEQ ID NO:748 and SEQ ID NO:1495
SEQ ID NO:749 and SEQ ID NO:1496
SEQ ID NO:750 and SEQ ID NO:1497
SEQ ID NO:751 and SEQ ID NO:1498
SEQ ID NO:752 and SEQ ID NO:1499
SEQ ID NO:753 and SEQ ID NO:1500
SEQ ID NO:754 and SEQ ID NO:1501
SEQ ID NO:755 and SEQ ID NO:1502
SEQ ID NO:756 and SEQ ID NO:1503
SEQ ID NO:757 and SEQ ID NO:1504
SEQ ID NO:758 and SEQ ID NO:1505
SEQ ID NO:759 and SEQ ID NO:1506
SEQ ID NO:760 and SEQ ID NO:1507
SEQ ID NO:761 and SEQ ID NO:1508
SEQ ID NO:762 and SEQ ID NO:1509
SEQ ID NO:763 and SEQ ID NO:1510
SEQ ID NO:764 and SEQ ID NO:1511
SEQ ID NO:765 and SEQ ID NO:1512
SEQ ID NO:766 and SEQ ID NO:1513
SEQ ID NO:767 and SEQ ID NO:1514
SEQ ID NO:768 and SEQ ID NO:1515
SEQ ID NO:769 and SEQ ID NO:1516
SEQ ID NO:770 and SEQ ID NO:1517
SEQ ID NO:771 and SEQ ID NO:1518
SEQ ID NO:772 and SEQ ID NO:1519
SEQ ID NO:773 and SEQ ID NO:1520
SEQ ID NO:774 and SEQ ID NO:1521
SEQ ID NO:775 and SEQ ID NO:1522
SEQ ID NO:776 and SEQ ID NO:1523
SEQ ID NO:777 and SEQ ID NO:1524
SEQ ID NO:778 and SEQ ID NO:1525
SEQ ID NO:779 and SEQ ID NO:1526
SEQ ID NO:780 and SEQ ID NO:1527
SEQ ID NO:781 and SEQ ID NO:1528
SEQ ID NO:782 and SEQ ID NO:1529
SEQ ID NO:783 and SEQ ID NO:1530
SEQ ID NO:784 and SEQ ID NO:1531
SEQ ID NO:785 and SEQ ID NO:1532
SEQ ID NO:786 and SEQ ID NO:1533
SEQ ID NO:787 and SEQ ID NO:1534
SEQ ID NO:788 and SEQ ID NO:1535
SEQ ID NO:789 and SEQ ID NO:1536
SEQ ID NO:790 and SEQ ID NO:1537
SEQ ID NO:791 and SEQ ID NO:1538
SEQ ID NO:792 and SEQ ID NO:1539
SEQ ID NO:793 and SEQ ID NO:1540
SEQ ID NO:794 and SEQ ID NO:1541
SEQ ID NO:795 and SEQ ID NO:1542
SEQ ID NO:797 and SEQ ID NO:1543
SEQ ID NO:798 and SEQ ID NO:1544

SEQ ID NO:799 and SEQ ID NO:1545
SEQ ID NO:800 and SEQ ID NO:1546
SEQ ID NO:801 and SEQ ID NO:1547
SEQ ID NO:802 and SEQ ID NO:1548
SEQ ID NO:803 and SEQ ID NO:1549
SEQ ID NO:804 and SEQ ID NO:1550
SEQ ID NO:805 and SEQ ID NO:1551
SEQ ID NO:806 and SEQ ID NO:1552
SEQ ID NO:807 and SEQ ID NO:1553
SEQ ID NO:809 and SEQ ID NO:1554
SEQ ID NO:810 and SEQ ID NO:1555
SEQ ID NO:811 and SEQ ID NO:1556
SEQ ID NO:812 and SEQ ID NO:1557
SEQ ID NO:813 and SEQ ID NO:1558
SEQ ID NO:814 and SEQ ID NO:1559
SEQ ID NO:816 and SEQ ID NO:1560
SEQ ID NO:817 and SEQ ID NO:1561
SEQ ID NO:818 and SEQ ID NO:1562
SEQ ID NO:819 and SEQ ID NO:1563
SEQ ID NO:820 and SEQ ID NO:1564
SEQ ID NO:822 and SEQ ID NO:1565
SEQ ID NO:823 and SEQ ID NO:1566
SEQ ID NO:824 and SEQ ID NO:1567
SEQ ID NO:825 and SEQ ID NO:1568
SEQ ID NO:826 and SEQ ID NO:1569
SEQ ID NO:827 and SEQ ID NO:1570
SEQ ID NO:828 and SEQ ID NO:1571
SEQ ID NO:829 and SEQ ID NO:1572, and
SEO ID NO:2545 and SEO ID NO:2546

4. A polynucleotide which can be synthesized with the primer set of claim 2 or 3.

5. A polynucleotide comprising a coding region in the polynucleotide of claim 4.

6. A substantially pure protein encoded by polynucleotide of claim 4.

7. A partial peptide of the protein of claim 6.

8. An isolated polynucleotide selected from the group consisting of

(a) a polynucleotide comprising a coding region of the nucleotide sequence set forth in any one of the following SEQ ID NOs:

SEQ ID NO:2547, SEQ ID NO:nnnn , SEQ ID NO:2551, SEQ ID NO:2553, SEQ ID NO:2555, SEQ ID NO:2557, SEQ ID NO:2558, SEQ ID NO:2560, SEQ ID NO:2562, SEQ ID NO:2564, SEQ ID NO:2566, SEQ ID NO:2568, SEQ ID NO:2570, SEQ ID NO:2572, SEQ ID NO:2574, SEQ ID NO:2576, SEQ ID NO:2578, SEQ ID NO:2580, SEQ ID NO:2581, SEQ ID NO:2583, SEQ ID NO:2584, SEQ ID NO:nnnn , SEQ ID NO:2588, SEQ ID NO:2590, SEQ ID NO:2592, SEQ ID NO:2594, SEQ ID NO:2596, SEQ ID NO:2598, SEQ ID NO:nnnn , SEQ ID NO:2602, SEQ ID NO:2604, SEQ ID NO:2606, SEQ ID NO:2607, SEQ ID NO:2609, SEQ ID NO:2611, SEQ ID NO:2613, SEQ ID NO:2615, SEQ ID NO:2617, SEQ ID NO:2619, SEQ ID NO:2621, SEQ ID NO:2623, SEQ ID NO:2625, SEQ ID NO:2627, SEQ ID NO:2629, SEQ ID NO:2630, SEQ ID NO:2632, SEQ ID NO:2634, SEQ ID NO:2636, SEQ ID NO:2638, SEQ ID NO:2640, SEQ ID NO:2642, SEQ ID NO:2644, SEQ ID NO:2646, SEQ ID NO:2648, SEQ ID NO:2650, SEQ ID NO:2652, SEQ ID NO:2654, SEQ ID NO:2656, SEQ ID NO:2658, SEQ ID NO:2660, SEQ ID NO:2662, SEQ ID NO:2664, SEQ ID NO:2666, SEQ ID NO:2668, SEQ ID NO:2670, SEQ ID NO:2672, SEQ ID NO:2674, SEQ ID NO:2676, SEQ ID NO:2678, SEQ ID NO:2680, SEQ ID NO:2681, SEQ ID NO:2683, SEQ ID NO:2685, SEQ ID NO:2686, SEQ ID NO:2688, SEQ ID NO:2690, SEQ ID NO:2692, SEQ ID NO:2694, SEQ ID NO:2696, SEQ ID NO:2698, SEQ ID NO:2700, SEQ ID NO:2702, SEQ ID NO:2704, SEQ ID NO:2706, SEQ ID NO:2708,

SEQ ID NO:2710, SEQ ID NO:2712, SEQ ID NO:2714, SEQ ID NO:2715, SEQ ID NO:2717, SEQ ID NO:2718, SEQ ID NO:2719, SEQ ID NO:2721, SEQ ID NO:2722, SEQ ID NO:2724, SEQ ID NO:2726, SEQ ID NO:2728, SEQ ID NO:2730, SEQ ID NO:2732, SEQ ID NO:2734 SEQ ID NO:2735, SEQ ID NO:2737, SEQ ID NO:2739, SEQ ID NO:nnnn , SEQ ID NO:2743, SEQ ID NO:2745, SEQ ID NO:2747, SEQ ID NO:2749, SEQ ID NO:2750, SEQ ID NO:2752, SEQ ID NO:2754, SEQ ID NO:2756, SEQ ID NO:2758, SEQ ID NO:2760, SEQ ID NO:2761, SEQ ID NO:nnnn , SEQ ID NO:2765, SEQ ID NO:2766, SEQ ID NO:2768, SEQ ID NO:2770, SEQ ID NO:2772, SEQ ID NO:2774, SEQ ID NO:2776, SEQ ID NO:2778, SEQ ID NO:2780, SEQ ID NO:2782, SEQ ID NO:2784, SEQ ID NO:2786, SEQ ID NO:2788, SEQ ID NO:2790, SEQ ID NO:2792, SEQ ID NO:2794, SEQ ID NO:2796, SEQ ID NO:2798, SEQ ID NO:2800, SEQ ID NO:2802, SEQ ID NO:2804, SEQ ID NO:2806, SEQ ID NO:2808, SEQ ID NO:2810, SEQ ID NO:2812, SEQ ID NO:2814, SEQ ID NO:2816, SEQ ID NO:2818, SEQ ID NO:2820, SEQ ID NO:2822, SEQ ID NO:2824, SEQ ID NO:2826, SEQ ID NO:2828, SEQ ID NO:2830, SEQ ID NO:2832, SEQ ID NO:2834, SEQ ID NO:2836, SEQ ID NO:2838, SEQ ID NO:2840, SEQ ID NO:2842, SEQ ID NO:2844, SEQ ID NO:2846, SEQ ID NO:2848, SEQ ID NO:2850, SEQ ID NO:2852, SEQ ID NO:2854, SEQ ID NO:2856, SEQ ID NO:2858, SEQ ID NO:2860, SEQ ID NO:2862, SEQ ID NO:2864, SEQ ID NO:2866, SEQ ID NO:2868, SEQ ID NO:2870, SEQ ID NO:2872, SEQ ID NO:2874, SEQ ID NO:2876, SEQ ID NO:2878, SEQ ID NO:2880, SEQ ID NO:2882, SEQ ID NO:2884, SEQ ID NO:2886, SEQ ID NO:2888, SEQ ID NO:2890, SEQ ID NO:2892, SEQ ID NO:2894, SEQ ID NO:2896, SEQ ID NO:2898, SEQ ID NO:2900, SEQ ID NO:2902, SEQ ID NO:2904, SEQ ID NO:2906, SEQ ID NO:2908, SEQ ID NO:2910, SEQ ID NO:2912, SEQ ID NO:2914, SEQ ID NO:2916, SEQ ID NO:2918, SEQ ID NO:2920, SEQ ID NO:2922, SEQ ID NO:2924, SEQ ID NO:2926, SEQ ID NO:2928, SEQ ID NO:2930 SEQ ID NO:2932, SEQ ID NO:2934, SEQ ID NO:2936, SEQ ID NO:2938, SEQ ID NO:2940, SEQ ID NO:2942, SEQ ID NO:2944, SEQ ID NO:2946, SEQ ID NO:2948, SEQ ID NO:2950, SEQ ID NO:2952, SEQ ID NO:2954, SEQ ID NO:2956, SEQ ID NO:2958, SEQ ID NO:2960, SEQ ID NO:2962, SEQ ID NO:nnnn , SEQ ID NO:2966, SEQ ID NO:2967, SEQ ID NO:2969, SEQ ID NO:2971, SEQ ID NO:2973, SEQ ID NO:2975, SEQ ID NO:2977, SEQ ID NO:2979, SEQ ID NO:2981, SEQ ID NO:2983, SEQ ID NO:2985, SEQ ID NO:2987, SEQ ID NO:2989, SEQ ID NO:2991, SEQ ID NO:2993, SEQ ID NO:2995, SEQ ID NO:2997, SEQ ID NO:2999, SEQ ID NO:3001, SEQ ID NO:3003, SEQ ID NO:3005, SEQ ID NO:3007, SEQ ID NO:3009, SEQ ID NO:3011, SEQ ID NO:3013, SEQ ID NO:3015, SEQ ID NO:3017, SEQ ID NO:3019, SEQ ID NO:3021, SEQ ID NO:3023, SEQ ID NO:3025, SEQ ID NO:3027, SEQ ID NO:3029, SEQ ID NO:3031, SEQ ID NO:3033, SEQ ID NO:3035, SEQ ID NO:3037, SEQ ID NO:3039, SEQ ID NO:nnnn , SEQ ID NO:nnnn , SEQ ID NO:3045, SEQ ID NO:3047, SEQ ID NO:3049, SEQ ID NO:3051, SEQ ID NO:3053, SEQ ID NO:3055, SEQ ID NO:3057, SEQ ID NO:3059, SEQ ID NO:3061, SEQ ID NO:3063, SEQ ID NO:3065, SEQ ID NO:3067, SEQ ID NO:3069, SEQ ID NO:3071, SEQ ID NO:3073, SEQ ID NO:3075, SEQ ID NO:3077, SEQ ID NO:3079, SEQ ID NO:3081, SEQ ID NO:3083, SEQ ID NO:3085, SEQ ID NO:3087, SEQ ID NO:3089, SEQ ID NO:3091, SEQ ID NO:3093, SEQ ID NO:3095, SEQ ID NO:3097, SEQ ID NO:3099, SEQ ID NO:3101, SEQ ID NO:3103, SEQ ID NO:3105, SEQ ID NO:3107, SEQ ID NO:3109, SEQ ID NO:3111, SEQ ID NO:3113, SEQ ID NO:3115, SEQ ID NO:3117, SEQ ID NO:3119, SEQ ID NO:3121, SEQ ID NO:3123, SEQ ID NO:3125, SEQ ID NO:3127, SEQ ID NO:3129 SEQ ID NO:3131, SEQ ID NO:3133, SEQ ID NO:3135, SEQ ID NO:3137, SEQ ID NO:3139, SEQ ID NO:3140, SEQ ID NO:3142, SEQ ID NO:3143, SEQ ID NO:3144, SEQ ID NO:3146, SEQ ID NO:3148, SEQ ID NO:3150, SEQ ID NO:3152, SEQ ID NO:3154, SEQ ID NO:3156, SEQ ID NO:3158, SEQ ID NO:nnnn , SEQ ID NO:3162, SEQ ID NO:3164, SEQ ID NO:3166, SEQ ID NO:3168, SEQ ID NO:3170, SEQ ID NO:3172, SEQ ID NO:3174, SEQ ID NO:3176, SEQ ID NO:3178, SEQ ID NO:3180, SEQ ID NO:3181, SEQ ID NO:3183, SEQ ID NO:3185, SEQ ID NO:3187, SEQ ID NO:3188, SEQ ID NO:3189, SEQ ID NO:3191, SEQ ID NO:3193, SEQ ID NO:3195, SEQ ID NO:3196, SEQ ID NO:3198, SEQ ID NO:3200, SEQ ID NO:3202, SEQ ID NO:3204, SEQ ID NO:3206, SEQ ID NO:3208, SEQ ID NO:3210, SEQ ID NO:3212,

SEQ ID NO:3214, SEQ ID NO:nnnn , SEQ ID NO:3218, SEQ ID NO:3219, SEQ ID NO:3221, SEQ ID NO:3222, SEQ ID NO:3224, SEQ ID NO:3226, SEQ ID NO:nnnn , SEQ ID NO:3230, SEQ ID NO:3232, SEQ ID NO:3234, SEQ ID NO:3236, SEQ ID NO:3238, SEQ ID NO:3240, SEQ ID NO:3242, SEQ ID NO:3244, SEQ ID NO:3246, SEQ ID NO:nnnn , SEQ ID NO:3250, SEQ ID NO:3252, SEQ ID NO:3254, SEQ ID NO:3256, SEQ ID NO:3258, SEQ ID NO:3260, SEQ ID NO:3262, SEQ ID NO:3264, SEQ ID NO:3266, SEQ ID NO:3268, SEQ ID NO:3270, SEQ ID NO:3272, SEQ ID NO:3274, SEQ ID NO:3276, SEQ ID NO:3278, SEQ ID NO:3280, SEQ ID NO:3282, SEQ ID NO:3284, SEQ ID NO:3286, SEQ ID NO:3288, SEQ ID NO:3290, SEQ ID NO:3292, SEQ ID NO:3294, SEQ ID NO:3296, SEQ ID NO:3298, SEQ ID NO:3300, SEQ ID NO:3301, SEQ ID NO:3303, SEQ ID NO:3305, SEQ ID NO:3307, SEQ ID NO:3309, SEQ ID NO:3311, SEQ ID NO:3312, SEQ ID NO:3314, SEQ ID NO:3316, SEQ ID NO:3318 SEQ ID NO:3320, SEQ ID NO:3322, SEQ ID NO:3324, SEQ ID NO:3326, SEQ ID NO:3328, SEQ ID NO:3330, SEQ ID NO:3332, SEQ ID NO:3334, SEQ ID NO:3336, SEQ ID NO:3338, SEQ ID NO:nnnn , SEQ ID NO:3342, SEQ ID NO:3344, SEQ ID NO:3346, SEQ ID NO:3348, SEQ ID NO:3350, SEQ ID NO:3352, SEQ ID NO:3354, SEQ ID NO:3356, SEQ ID NO:3358, SEQ ID NO:3360, SEQ ID NO:3362, SEQ ID NO:3364, SEQ ID NO:3366, SEQ ID NO:3368, SEQ ID NO:3370, SEQ ID NO:3372, SEQ ID NO:3374, SEQ ID NO:3376, SEQ ID NO:3378, SEQ ID NO:3380, SEQ ID NO:3382, SEQ ID NO:3384, SEQ ID NO:3386, SEQ ID NO:3388, SEQ ID NO:3390, SEQ ID NO:3392, SEQ ID NO:3394, SEQ ID NO:3396, SEQ ID NO:3398, SEQ ID NO:3400, SEQ ID NO:3402, SEQ ID NO:3404, SEQ ID NO:nnnn , SEQ ID NO:3408, SEQ ID NO:3410, SEQ ID NO:3411, SEQ ID NO:3413, SEQ ID NO:3415, SEQ ID NO:3417, SEQ ID NO:3419, SEQ ID NO:nnnn , SEQ ID NO:3423, SEQ ID NO:3425, SEQ ID NO:3427, SEQ ID NO:3429, SEQ ID NO:3431, SEQ ID NO:3433, SEQ ID NO:3435, SEQ ID NO:3437, SEQ ID NO:3439, SEQ ID NO:3441, SEQ ID NO:3443, SEQ ID NO:3445, SEQ ID NO:3447, SEQ ID NO:3449, SEQ ID NO:3451, SEQ ID NO:3453, SEQ ID NO:3455, SEQ ID NO:3457, SEQ ID NO:3459, SEQ ID NO:3461, SEQ ID NO:3463, SEQ ID NO:3465, SEQ ID NO:3467, SEQ ID NO:3469, SEQ ID NO:3471, SEQ ID NO:3473, SEQ ID NO:3475, SEQ ID NO:3477, SEQ ID NO:3479, SEQ ID NO:3481, SEQ ID NO:3482, SEQ ID NO:3484, SEQ ID NO:3486, SEQ ID NO:3488, SEQ ID NO:3490, SEQ ID NO:3492, SEQ ID NO:3494, SEQ ID NO:3496, SEQ ID NO:3498, SEQ ID NO:3500, SEQ ID NO:3502, SEQ ID NO:3504, SEQ ID NO:3506, SEQ ID NO:3508, SEQ ID NO:3509, SEQ ID NO:3511, SEQ ID NO:3513, SEQ ID NO:3515 SEQ ID NO:3517, SEQ ID NO:3519, SEQ ID NO:3521, SEQ ID NO:3523, SEQ ID NO:3525, SEQ ID NO:3527, SEQ ID NO:3529, SEQ ID NO:3531, SEQ ID NO:3533, SEQ ID NO:3535, SEQ ID NO:3537, SEQ ID NO:3539, SEQ ID NO:3541, SEQ ID NO:3543, SEQ ID NO:3545, SEQ ID NO:3547, SEQ ID NO:3549, SEQ ID NO:3551, SEQ ID NO:3553, SEQ ID NO:3555, SEQ ID NO:3557, SEQ ID NO:3559, SEQ ID NO:3561, SEQ ID NO:3563, SEQ ID NO:nnnn , SEQ ID NO:3567, SEQ ID NO:3569, SEQ ID NO:3571, SEQ ID NO:3573, SEQ ID NO:3575, SEQ ID NO:3577, SEQ ID NO:3579, SEQ ID NO:3581, SEQ ID NO:3583, SEQ ID NO:3585, SEQ ID NO:3587, SEQ ID NO:3589 , SEQ ID NO:3595, SEQ ID NO:3597, SEQ ID NO:3599, SEQ ID NO:3601, SEQ ID NO:3603, SEQ ID NO:3605, SEQ ID NO:3607, SEQ ID NO:3609, SEQ ID NO:3610, SEQ ID NO:3611, SEQ ID NO:3612, SEQ ID NO:3613, SEQ ID NO:3615, SEQ ID NO:3617, SEQ ID NO:3619, SEQ ID NO:3621, SEQ ID NO:3622, SEQ ID NO:3624, SEQ ID NO:3626, SEQ ID NO:3627, SEQ ID NO:3628, SEQ ID NO:3629, SEQ ID NO:3631, SEQ ID NO:3633, SEQ ID NO:3634, SEQ ID NO:3636, SEQ ID NO:3638, SEQ ID NO:3640, SEQ ID NO:3642, SEQ ID NO:3644, SEQ ID NO:3645, SEQ ID NO:3647, SEQ ID NO:3649, SEQ ID NO:3651, SEQ ID NO:3653, SEQ ID NO:3655, SEQ ID NO:3657, SEQ ID NO:3658, SEQ ID NO:3660, SEQ ID NO:3662, SEQ ID NO:3664, SEQ ID NO:3666, SEQ ID NO:3668, SEQ ID NO:3670, SEQ ID NO:3672, SEQ ID NO:3674, SEQ ID NO:3676, SEQ ID NO:3678, SEQ ID NO:3680, SEQ ID NO:3682, SEQ ID NO:3683, SEQ ID NO:3685, SEQ ID NO:3686, SEQ ID NO:3688, SEQ ID NO:3690, SEQ ID NO:3692, SEQ ID NO:3694, SEQ ID NO:3696, SEQ ID NO:3698, SEQ ID NO:3700, SEQ ID NO:3702, SEQ ID NO:3704, SEQ ID NO:3706, SEQ ID NO:3708, SEQ ID NO:3710, SEQ ID NO:3712, SEQ ID NO:3714, SEQ ID NO:3716,

SEQ ID NO:3718, SEQ ID NO:3720, SEQ ID NO:3722, SEQ ID NO:3724, SEQ ID NO:3726, SEQ ID NO:3728, SEQ ID NO:3730, SEQ ID NO:3732, SEQ ID NO:3734, SEQ ID NO:3736, SEQ ID NO:3738, SEQ ID NO:3740, SEQ ID NO:3741, SEQ ID NO:3743, SEQ ID NO:3745, SEQ ID NO:3747, SEQ ID NO:3749, SEQ ID NO:3751, SEQ ID NO:3753, SEQ ID NO:3755, SEQ ID NO:3756, SEQ ID NO:3758, SEQ ID NO:3760, SEQ ID NO:3761, SEQ ID NO:3763, SEQ ID NO:3765, SEQ ID NO:3767, SEQ ID NO:3769, SEQ ID NO:3771, SEQ ID NO:3773, SEQ ID NO:3775, SEQ ID NO:3777
SEQ ID NO:3779, SEQ ID NO:3781, SEQ ID NO:3783, SEQ ID NO:3785, SEQ ID NO:3787, SEQ ID NO:3789, SEQ ID NO:3791, SEQ ID NO:3793, SEQ ID NO:3795, SEQ ID NO:3797, SEQ ID NO:3799, SEQ ID NO:3801, SEQ ID NO:3803, SEQ ID NO:3805, SEQ ID NO:3806, SEQ ID NO:3808, SEQ ID NO:3810, SEQ ID NO:3812, SEQ ID NO:3814, SEQ ID NO:3816, SEQ ID NO:3818, SEQ ID NO:3820, SEQ ID NO:3822, SEQ ID NO:3824, SEQ ID NO:3826, SEQ ID NO:3828, SEQ ID NO:3830, SEQ ID NO:3832, SEQ ID NO:3834, SEQ ID NO:3836, SEQ ID NO:3838, SEQ ID NO:3840, SEQ ID NO:3842, SEQ ID NO:3844, SEQ ID NO:3846, SEQ ID NO:3848, SEQ ID NO:3852, SEQ ID NO:3854, SEQ ID NO:3856, SEQ ID NO:3858, SEQ ID NO:3860, SEQ ID NO:3862, SEQ ID NO:3864, SEQ ID NO:3866, SEQ ID NO:3868, SEQ ID NO:3870, SEQ ID NO:3872, SEQ ID NO:3874, SEQ ID NO:3876, SEQ ID NO:3878, SEQ ID NO:3880, SEQ ID NO:3882, SEQ ID NO:3884, SEQ ID NO:3886, SEQ ID NO:3888, SEQ ID NO:3890, SEQ ID NO:3892, SEQ ID NO:3894, SEQ ID NO:3896, SEQ ID NO:3898, SEQ ID NO:3900, SEQ ID NO:3902, SEQ ID NO:3904, SEQ ID NO:3906, SEQ ID NO:3908, SEQ ID NO:3910, SEQ ID NO:3912, SEQ ID NO:3914, SEQ ID NO:3916, SEQ ID NO:3918, SEQ ID NO:3920, SEQ ID NO:3922, SEQ ID NO:3924, SEQ ID NO:3926, SEQ ID NO:3928, SEQ ID NO:3930, SEQ ID NO:3932, SEQ ID NO:3934, SEQ ID NO:3936, SEQ ID NO:3938, SEQ ID NO:3940, SEQ ID NO:3942, SEQ ID NO:3944, SEQ ID NO:3946, SEQ ID NO:3948, SEQ ID NO:3950, SEQ ID NO:3952, SEQ ID NO:3954, SEQ ID NO:3956, SEQ ID NO:3958, SEQ ID NO:3960, SEQ ID NO:3962, SEQ ID NO:3964, SEQ ID NO:3965, SEQ ID NO:3966, SEQ ID NO:3968, SEQ ID NO:3970, SEQ ID NO:3972, SEQ ID NO:3974, SEQ ID NO:3976
SEQ ID NO:3977, SEQ ID NO:3979, SEQ ID NO:3981, SEQ ID NO:3983, SEQ ID NO:3985, SEQ ID NO:3987, SEQ ID NO:3988, SEQ ID NO:3990, SEQ ID NO:3992, SEQ ID NO:3994, SEQ ID NO:3996, SEQ ID NO:3998, SEQ ID NO:4000, SEQ ID NO:4001, SEQ ID NO:4003, SEQ ID NO:4004, SEQ ID NO:4006, SEQ ID NO:4008, SEQ ID NO:4010, SEQ ID NO:4012, SEQ ID NO:4016, SEQ ID NO:4018, SEQ ID NO:4020, SEQ ID NO:4022, SEQ ID NO:4024, SEQ ID NO:4026, SEQ ID NO:4028, SEQ ID NO:4030, SEQ ID NO:4032, SEQ ID NO:4034, SEQ ID NO:4036, SEQ ID NO:4038, SEQ ID NO:4040, SEQ ID NO:4042, SEQ ID NO:4044, SEQ ID NO:4046, SEQ ID NO:4048, SEQ ID NO:4050, SEQ ID NO:4052, SEQ ID NO:4054, SEQ ID NO:4056, SEQ ID NO:4058, SEQ ID NO:4060, SEQ ID NO:4062, SEQ ID NO:4064, SEQ ID NO:4066, SEQ ID NO:4068, SEQ ID NO:4070, SEQ ID NO:4072, SEQ ID NO:4074, SEQ ID NO:4076
SEQ ID NO:4078, SEQ ID NO:4080, SEQ ID NO:4082, SEQ ID NO:4084, SEQ ID NO:4086, SEQ ID NO:4088, SEQ ID NO:4090, SEQ ID NO:4092, SEQ ID NO:4094, SEQ ID NO:4096, SEQ ID NO:4098, SEQ ID NO:4099, SEQ ID NO:4101, SEQ ID NO:4103, SEQ ID NO:4105, SEQ ID NO:4107, SEQ ID NO:4109, SEQ ID NO:4111, SEQ ID NO:4113, SEQ ID NO:4115, SEQ ID NO:4117, SEQ ID NO:4119, SEQ ID NO:4121, SEQ ID NO:4123, SEQ ID NO:4125, SEQ ID NO:4127, SEQ ID NO:4129, SEQ ID NO:4131, SEQ ID NO:4132, SEQ ID NO:4134, SEQ ID NO:4136, SEQ ID NO:4137, SEQ ID NO:4139, SEQ ID NO:4140, SEQ ID NO:4142, SEQ ID NO:4144, SEQ ID NO:4146, SEQ ID NO:4148, SEQ ID NO:4150, SEQ ID NO:4152, SEQ ID NO:4154, SEQ ID NO:4155, SEQ ID NO:4157, SEQ ID NO:4159, SEQ ID NO:4161, SEQ ID NO:4163, SEQ ID NO:4165
SEQ ID NO:4167, SEQ ID NO:4168, SEQ ID NO:4170, SEQ ID NO:4172, SEQ ID NO:4174, SEQ ID NO:4176, and SEQ ID NO:4178

(b) a polynucleotide comprising a nucleotide sequence encoding a protein comprising the amino acid sequence

set forth in any one of the following SEQ ID NOs:

SEQ ID NO:2548, SEQ ID NO:nnnn , SEQ ID NO:2552, SEQ ID NO:2554, SEQ ID NO:2556, SEQ ID NO:2559, SEQ ID NO:2561, SEQ ID NO:2563, SEQ ID NO:2565, SEQ ID NO:2567, SEQ ID NO:2569, SEQ ID NO:2571, SEQ ID NO:2573, SEQ ID NO:2575, SEQ ID NO:2577, SEQ ID NO:2579, SEQ ID NO:2582, SEQ ID NO:2585, SEQ ID NO:nnnn , SEQ ID NO:2589, SEQ ID NO:2591, SEQ ID NO:2593, SEQ ID NO:2595, SEQ ID NO:2597, SEQ ID NO:2599, SEQ ID NO:nnnn , SEQ ID NO:2603, SEQ ID NO:2605, SEQ ID NO:2608, SEQ ID NO:2610, SEQ ID NO:2612, SEQ ID NO:2614, SEQ ID NO:2616, SEQ ID NO:2618, SEQ ID NO:2620, SEQ ID NO:2622, SEQ ID NO:2624, SEQ ID NO:2626, SEQ ID NO:2628, SEQ ID NO:2631, SEQ ID NO:2633, SEQ ID NO:2635, SEQ ID NO:2637, SEQ ID NO:2639, SEQ ID NO:2641, SEQ ID NO:2643, SEQ ID NO:2645, SEQ ID NO:2647, SEQ ID NO:2649, SEQ ID NO:2651, SEQ ID NO:2653, SEQ ID NO:2655, SEQ ID NO:2657, SEQ ID NO:2659, SEQ ID NO:2661, SEQ ID NO:2663, SEQ ID NO:2665, SEQ ID NO:2667, SEQ ID NO:2669, SEQ ID NO:2671, SEQ ID NO:2673, SEQ ID NO:2675, SEQ ID NO:2677, SEQ ID NO:2679, SEQ ID NO:2682, SEQ ID NO:2684, SEQ ID NO:2687, SEQ ID NO:2689, SEQ ID NO:2691, SEQ ID NO:2693, SEQ ID NO:2695, SEQ ID NO:2697, SEQ ID NO:2699, SEQ ID NO:2701, SEQ ID NO:2703, SEQ ID NO:2705, SEQ ID NO:2707, SEQ ID NO:2709, SEQ ID NO:2711, SEQ ID NO:2713, SEQ ID NO:2716, SEQ ID NO:2720, SEQ ID NO:2723, SEQ ID NO:2725, SEQ ID NO:2727, SEQ ID NO:2729, SEQ ID NO:2731, SEQ ID NO:2733, SEQ ID NO:2736, SEQ ID NO:2738, SEQ ID NO:2740, SEQ ID NO:nnnn , SEQ ID NO:2744, SEQ ID NO:2746, SEQ ID NO:2748, SEQ ID NO:2751, SEQ ID NO:2753, SEQ ID NO:2755, SEQ ID NO:2757, SEQ ID NO:2759 SEQ ID NO:2762, SEQ ID NO:nnnn , SEQ ID NO:2767, SEQ ID NO:2769, SEQ ID NO:2771, SEQ ID NO:2773, SEQ ID NO:2775, SEQ ID NO:2777, SEQ ID NO:2779, SEQ ID NO:2781, SEQ ID NO:2783, SEQ ID NO:2785, SEQ ID NO:2787, SEQ ID NO:2789, SEQ ID NO:2791, SEQ ID NO:2793, SEQ ID NO:2795, SEQ ID NO:2797, SEQ ID NO:2799, SEQ ID NO:2801, SEQ ID NO:2803, SEQ ID NO:2805, SEQ ID NO:2807, SEQ ID NO:2809, SEQ ID NO:2811, SEQ ID NO:2813, SEQ ID NO:2815, SEQ ID NO:2817, SEQ ID NO:2819, SEQ ID NO:2821, SEQ ID NO:2823, SEQ ID NO:2825, SEQ ID NO:2827, SEQ ID NO:2829, SEQ ID NO:2831, SEQ ID NO:2833, SEQ ID NO:2835, SEQ ID NO:2837, SEQ ID NO:2839, SEQ ID NO:2841, SEQ ID NO:2843, SEQ ID NO:2845, SEQ ID NO:2847, SEQ ID NO:2849, SEQ ID NO:2851, SEQ ID NO:2853, SEQ ID NO:2855, SEQ ID NO:2857, SEQ ID NO:2859, SEQ ID NO:2861, SEQ ID NO:2863, SEQ ID NO:2865, SEQ ID NO:2867, SEQ ID NO:2869, SEQ ID NO:2871, SEQ ID NO:2873, SEQ ID NO:2875, SEQ ID NO:2877, SEQ ID NO:2879, SEQ ID NO:2881, SEQ ID NO:2883, SEQ ID NO:2885, SEQ ID NO:2887, SEQ ID NO:2889, SEQ ID NO:2891, SEQ ID NO:2893, SEQ ID NO:2895, SEQ ID NO:2897, SEQ ID NO:2899, SEQ ID NO:2901, SEQ ID NO:2903, SEQ ID NO:2905, SEQ ID NO:2907, SEQ ID NO:2909, SEQ ID NO:2911, SEQ ID NO:2913, SEQ ID NO:2915, SEQ ID NO:2917, SEQ ID NO:2919, SEQ ID NO:2921, SEQ ID NO:2923, SEQ ID NO:2925, SEQ ID NO:2927, SEQ ID NO:2929, SEQ ID NO:2931, SEQ ID NO:2933, SEQ ID NO:2935, SEQ ID NO:2937, SEQ ID NO:2939, SEQ ID NO:2941, SEQ ID NO:2943, SEQ ID NO:2945, SEQ ID NO:2947, SEQ ID NO:2949, SEQ ID NO:2951, SEQ ID NO:2953, SEQ ID NO:2955, SEQ ID NO:2957, SEQ ID NO:2959, SEQ ID NO:2961 SEQ ID NO:2963, SEQ ID NO:nnnn , SEQ ID NO:2968, SEQ ID NO:2970, SEQ ID NO:2972, SEQ ID NO:2974, SEQ ID NO:2976, SEQ ID NO:2978, SEQ ID NO:2980, SEQ ID NO:2982, SEQ ID NO:2984, SEQ ID NO:2986, SEQ ID NO:2988, SEQ ID NO:2990, SEQ ID NO:2992, SEQ ID NO:2994, SEQ ID NO:2996, SEQ ID NO:2998, SEQ ID NO:3000, SEQ ID NO:3002, SEQ ID NO:3004, SEQ ID NO:3006, SEQ ID NO:3008, SEQ ID NO:3010, SEQ ID NO:3012, SEQ ID NO:3014, SEQ ID NO:3016, SEQ ID NO:3018, SEQ ID NO:3020, SEQ ID NO:3022, SEQ ID NO:3024, SEQ ID NO:3026, SEQ ID NO:3028, SEQ ID NO:3030, SEQ ID NO:3032, SEQ ID NO:3034, SEQ ID NO:3036, SEQ ID NO:3038, SEQ ID NO:3040, SEQ ID NO:nnnn , SEQ ID NO:nnnn , SEQ ID NO:3046, SEQ ID NO:3048, SEQ ID NO:3050, SEQ ID NO:3052, SEQ ID NO:3054, SEQ ID NO:3056, SEQ ID NO:3058, SEQ ID NO:3060, SEQ ID NO:3062, SEQ ID NO:3064, SEQ ID NO:3066, SEQ ID NO:3068, SEQ ID NO:3070, SEQ ID NO:3072, SEQ ID NO:3074, SEQ ID NO:3076, SEQ ID NO:3078, SEQ ID NO:3080, SEQ ID NO:3082,

SEQ ID NO:3084, SEQ ID NO:3086, SEQ ID NO:3088, SEQ ID NO:3090, SEQ ID NO:3092, SEQ ID NO:3094, SEQ ID NO:3096, SEQ ID NO:3098, SEQ ID NO:3100, SEQ ID NO:3102, SEQ ID NO:3104, SEQ ID NO:3106, SEQ ID NO:3108, SEQ ID NO:3110, SEQ ID NO:3112, SEQ ID NO:3114, SEQ ID NO:3116, SEQ ID NO:3118, SEQ ID NO:3120, SEQ ID NO:3122, SEQ ID NO:3124, SEQ ID NO:3126, SEQ ID NO:3128, SEQ ID NO:3130, SEQ ID NO:3132, SEQ ID NO:3134, SEQ ID NO:3136, SEQ ID NO:3138, SEQ ID NO:3141, SEQ ID NO:3145, SEQ ID NO:3147, SEQ ID NO:3149, SEQ ID NO:3151, SEQ ID NO:3153, SEQ ID NO:3155, SEQ ID NO:3157, SEQ ID NO:3159, SEQ ID NO:nnnn , SEQ ID NO:3163, SEQ ID NO:3165 SEQ ID NO:3167, SEQ ID NO:3169, SEQ ID NO:3171, SEQ ID NO:3173, SEQ ID NO:3175, SEQ ID NO:3177, SEQ ID NO:3179, SEQ ID NO:3182, SEQ ID NO:3184, SEQ ID NO:3186, SEQ ID NO:3190, SEQ ID NO:3192, SEQ ID NO:3194, SEQ ID NO:3197, SEQ ID NO:3199, SEQ ID NO:3201, SEQ ID NO:3203, SEQ ID NO:3205, SEQ ID NO:3207, SEQ ID NO:3209, SEQ ID NO:3211, SEQ ID NO:3213, SEQ ID NO:3215, SEQ ID NO:nnnn , SEQ ID NO:3220, SEQ ID NO:3223, SEQ ID NO:3225, SEQ ID NO:3227, SEQ ID NO:nnnn , SEQ ID NO:3231, SEQ ID NO:3233, SEQ ID NO:3235, SEQ ID NO:3237, SEQ ID NO:3239, SEQ ID NO:3241, SEQ ID NO:3243, SEQ ID NO:3245, SEQ ID NO:3247, SEQ ID NO:nnnn , SEQ ID NO:3251, SEQ ID NO:3253, SEQ ID NO:3255, SEQ ID NO:3257, SEQ ID NO:3259, SEQ ID NO:3261, SEQ ID NO:3263, SEQ ID NO:3265, SEQ ID NO:3267, SEQ ID NO:3269, SEQ ID NO:3271, SEQ ID NO:3273, SEQ ID NO:3275, SEQ ID NO:3277, SEQ ID NO:3279, SEQ ID NO:3281, SEQ ID NO:3283, SEQ ID NO:3285, SEQ ID NO:3287, SEQ ID NO:3289, SEQ ID NO:3291, SEQ ID NO:3293, SEQ ID NO:3295, SEQ ID NO:3297, SEQ ID NO:3299, SEQ ID NO:3302, SEQ ID NO:3304, SEQ ID NO:3306, SEQ ID NO:3308, SEQ ID NO:3310, SEQ ID NO:3313, SEQ ID NO:3315, SEQ ID NO:3317, SEQ ID NO:3319, SEQ ID NO:3321, SEQ ID NO:3323, SEQ ID NO:3325, SEQ ID NO:3327, SEQ ID NO:3329, SEQ ID NO:3331, SEQ ID NO:3333, SEQ ID NO:3335, SEQ ID NO:3337, SEQ ID NO:3339, SEQ ID NO:nnnn , SEQ ID NO:3343, SEQ ID NO:3345, SEQ ID NO:3347, SEQ ID NO:3349, SEQ ID NO:3351, SEQ ID NO:3353, SEQ ID NO:3355, SEQ ID NO:3357, SEQ ID NO:3359, SEQ ID NO:3361, SEQ ID NO:3363, SEQ ID NO:3365, SEQ ID NO:3367, SEQ ID NO:3369, SEQ ID NO:3371, SEQ ID NO:3373 SEQ ID NO:3375, SEQ ID NO:3377, SEQ ID NO:3379, SEQ ID NO:3381, SEQ ID NO:3383, SEQ ID NO:3385, SEQ ID NO:3387, SEQ ID NO:3389, SEQ ID NO:3391, SEQ ID NO:3393, SEQ ID NO:3395, SEQ ID NO:3397, SEQ ID NO:3399, SEQ ID NO:3401, SEQ ID NO:3403, SEQ ID NO:3405, SEQ ID NO:nnnn , SEQ ID NO:3409, SEQ ID NO:3412, SEQ ID NO:3414, SEQ ID NO:3416, SEQ ID NO:3418, SEQ ID NO:3420, SEQ ID NO:nnnn , SEQ ID NO:3424, SEQ ID NO:3426, SEQ ID NO:3428, SEQ ID NO:3430, SEQ ID NO:3432, SEQ ID NO:3434, SEQ ID NO:3436, SEQ ID NO:3438, SEQ ID NO:3440, SEQ ID NO:3442, SEQ ID NO:3444, SEQ ID NO:3446, SEQ ID NO:3448, SEQ ID NO:3450, SEQ ID NO:3452, SEQ ID NO:3454, SEQ ID NO:3456, SEQ ID NO:3458, SEQ ID NO:3460, SEQ ID NO:3462, SEQ ID NO:3464, SEQ ID NO:3466, SEQ ID NO:3468, SEQ ID NO:3470, SEQ ID NO:3472, SEQ ID NO:3474, SEQ ID NO:3476, SEQ ID NO:3478, SEQ ID NO:3480, SEQ ID NO:3483, SEQ ID NO:3485, SEQ ID NO:3487, SEQ ID NO:3489, SEQ ID NO:3491, SEQ ID NO:3493, SEQ ID NO:3495, SEQ ID NO:3497, SEQ ID NO:3499, SEQ ID NO:3501, SEQ ID NO:3503, SEQ ID NO:3505, SEQ ID NO:3507, SEQ ID NO:3510, SEQ ID NO:3512, SEQ ID NO:3514, SEQ ID NO:3516, SEQ ID NO:3518, SEQ ID NO:3520, SEQ ID NO:3522, SEQ ID NO:3524, SEQ ID NO:3526, SEQ ID NO:3528, SEQ ID NO:3530, SEQ ID NO:3532, SEQ ID NO:3534, SEQ ID NO:3536, SEQ ID NO:3538, SEQ ID NO:3540, SEQ ID NO:3542, SEQ ID NO:3544, SEQ ID NO:3546, SEQ ID NO:3548, SEQ ID NO:3550, SEQ ID NO:3552, SEQ ID NO:3554, SEQ ID NO:3556, SEQ ID NO:3558, SEQ ID NO:3560, SEQ ID NO:3562, SEQ ID NO:3564, SEQ ID NO:nnnn , SEQ ID NO:3568, SEQ ID NO:3570, SEQ ID NO:3572, SEQ ID NO:3574, SEQ ID NO:3576 SEQ ID NO:3578, SEQ ID NO:3580, SEQ ID NO:3582, SEQ ID NO:3584, SEQ ID NO:3586, SEQ ID NO:3588, SEQ ID NO:3590

SEQ ID NO:3596, SEQ ID NO:3598, SEQ ID NO:3600, SEQ ID NO:3602, SEQ ID NO:3604, SEQ ID NO:3606, SEQ ID NO:3608, SEQ ID NO:3614, SEQ ID NO:3616, SEQ ID NO:3618,

SEQ ID NO:3620, SEQ ID NO:3623, SEQ ID NO:3625, SEQ ID NO:3630, SEQ ID NO:3632, SEQ ID NO:3635, SEQ ID NO:3637, SEQ ID NO:3639, SEQ ID NO:3641, SEQ ID NO:3643, SEQ ID NO:3646, SEQ ID NO:3648, SEQ ID NO:3650, SEQ ID NO:3652, SEQ ID NO:3654, SEQ ID NO:3656, SEQ ID NO:3659, SEQ ID NO:3661, SEQ ID NO:3663, SEQ ID NO:3665, SEQ ID NO:3667, SEQ ID NO:3669, SEQ ID NO:3671, SEQ ID NO:3673, SEQ ID NO:3675, SEQ ID NO:3677, SEQ ID NO:3679, SEQ ID NO:3681, SEQ ID NO:3684, SEQ ID NO:3687, SEQ ID NO:3689, SEQ ID NO:3691, SEQ ID NO:3693, SEQ ID NO:3695, SEQ ID NO:3697, SEQ ID NO:3699, SEQ ID NO:3701, SEQ ID NO:3703, SEQ ID NO:3705, SEQ ID NO:3707, SEQ ID NO:3709, SEQ ID NO:3711, SEQ ID NO:3713, SEQ ID NO:3715, SEQ ID NO:3717, SEQ ID NO:3719, SEQ ID NO:3721, SEQ ID NO:3723, SEQ ID NO:3725, SEQ ID NO:3727, SEQ ID NO:3729, SEQ ID NO:3731, SEQ ID NO:3733, SEQ ID NO:3735, SEQ ID NO:3737, SEQ ID NO:3739, SEQ ID NO:3742, SEQ ID NO:3744, SEQ ID NO:3746, SEQ ID NO:3748, SEQ ID NO:3750, SEQ ID NO:3752, SEQ ID NO:3754, SEQ ID NO:3757, SEQ ID NO:3759, SEQ ID NO:3762, SEQ ID NO:3764, SEQ ID NO:3766, SEQ ID NO:3768, SEQ ID NO:3770, SEQ ID NO:3772, SEQ ID NO:3774, SEQ ID NO:3776, SEQ ID NO:3778, SEQ ID NO:3780, SEQ ID NO:3782, SEQ ID NO:3784, SEQ ID NO:3786, SEQ ID NO:3788, SEQ ID NO:3790, SEQ ID NO:3792, SEQ ID NO:3794, SEQ ID NO:3796, SEQ ID NO:3798, SEQ ID NO:3800, SEQ ID NO:3802, SEQ ID NO:3804, SEQ ID NO:3807, SEQ ID NO:3809, SEQ ID NO:3811 SEQ ID NO:3813, SEQ ID NO:3815, SEQ ID NO:3817, SEQ ID NO:3819, SEQ ID NO:3821, SEQ ID NO:3823, SEQ ID NO:3825, SEQ ID NO:3827, SEQ ID NO:3829, SEQ ID NO:3831, SEQ ID NO:3833, SEQ ID NO:3835, SEQ ID NO:3837, SEQ ID NO:3839, SEQ ID NO:3841, SEQ ID NO:3843, SEQ ID NO:3845, SEQ ID NO:3847, SEQ ID NO:3849, SEQ ID NO:3853, SEQ ID NO:3855, SEQ ID NO:3857, SEQ ID NO:3859, SEQ ID NO:3861, SEQ ID NO:3863, SEQ ID NO:3865, SEQ ID NO:3867, SEQ ID NO:3869, SEQ ID NO:3871, SEQ ID NO:3873, SEQ ID NO:3875, SEQ ID NO:3877, SEQ ID NO:3879, SEQ ID NO:3881, SEQ ID NO:3883, SEQ ID NO:3885, SEQ ID NO:3887, SEQ ID NO:3889, SEQ ID NO:3891, SEQ ID NO:3893, SEQ ID NO:3895, SEQ ID NO:3897, SEQ ID NO:3899, SEQ ID NO:3901, SEQ ID NO:3903, SEQ ID NO:3905, SEQ ID NO:3907, SEQ ID NO:3909, SEQ ID NO:3911, SEQ ID NO:3913, SEQ ID NO:3915, SEQ ID NO:3917, SEQ ID NO:3919, SEQ ID NO:3921, SEQ ID NO:3923, SEQ ID NO:3925, SEQ ID NO:3927, SEQ ID NO:3929, SEQ ID NO:3931, SEQ ID NO:3933, SEQ ID NO:3935, SEQ ID NO:3937, SEQ ID NO:3939, SEQ ID NO:3941, SEQ ID NO:3943, SEQ ID NO:3945, SEQ ID NO:3947, SEQ ID NO:3949, SEQ ID NO:3951, SEQ ID NO:3953, SEQ ID NO:3955, SEQ ID NO:3957, SEQ ID NO:3959, SEQ ID NO:3961, SEQ ID NO:3963, SEQ ID NO:3967, SEQ ID NO:3969, SEQ ID NO:3971, SEQ ID NO:3973, SEQ ID NO:3975, SEQ ID NO:3978, SEQ ID NO:3980, SEQ ID NO:3982, SEQ ID NO:3984, SEQ ID NO:3986, SEQ ID NO:3989, SEQ ID NO:3991, SEQ ID NO:3993, SEQ ID NO:3995, SEQ ID NO:3997, SEQ ID NO:3999, SEQ ID NO:4002, SEQ ID NO:4005, SEQ ID NO:4007, SEQ ID NO:4009, SEQ ID NO:4011, SEQ ID NO:4013, SEQ ID NO:4017, SEQ ID NO:4019, SEQ ID NO:4021 SEQ ID NO:4023, SEQ ID NO:4025, SEQ ID NO:4027, SEQ ID NO:4029, SEQ ID NO:4031, SEQ ID NO:4033, SEQ ID NO:4035, SEQ ID NO:4037, SEQ ID NO:4039, SEQ ID NO:4041, SEQ ID NO:4043, SEQ ID NO:4045, SEQ ID NO:4047, SEQ ID NO:4049, SEQ ID NO:4051, SEQ ID NO:4053, SEQ ID NO:4055, SEQ ID NO:4057, SEQ ID NO:4059, SEQ ID NO:4061, SEQ ID NO:4063, SEQ ID NO:4065, SEQ ID NO:4067, SEQ ID NO:4069, SEQ ID NO:4071, SEQ ID NO:4073, SEQ ID NO:4075, SEQ ID NO:4077
SEQ ID NO:4079, SEQ ID NO:4081, SEQ ID NO:4083, SEQ ID NO:4085, SEQ ID NO:4087, SEQ ID NO:4089, SEQ ID NO:4091, SEQ ID NO:4093, SEQ ID NO:4095, SEQ ID NO:4097, SEQ ID NO:4100, SEQ ID NO:4102, SEQ ID NO:4104, SEQ ID NO:4106, SEQ ID NO:4108, SEQ ID NO:4110, SEQ ID NO:4112, SEQ ID NO:4114, SEQ ID NO:4116, SEQ ID NO:4118, SEQ ID NO:4120, SEQ ID NO:4122, SEQ ID NO:4124, SEQ ID NO:4126, SEQ ID NO:4128, SEQ ID NO:4130, SEQ ID NO:4133, SEQ ID NO:4135, SEQ ID NO:4138, SEQ ID NO:4141, SEQ ID NO:4143, SEQ ID NO:4145, SEQ ID NO:4147, SEQ ID NO:4149, SEQ ID NO:4151, SEQ ID NO:4153, SEQ ID NO:4156, SEQ ID NO:4158, SEQ ID NO:4160, SEQ ID NO:4162,

SEQ ID NO:4164, SEQ ID NO:4166
SEQ ID NO:4169, SEQ ID NO:4171, SEQ ID NO:4173, SEQ ID NO:4175, SEQ ID NO:4177, and SEQ ID NO:4179

(c) a polynucleotide comprising a nucleotide sequence encoding a protein comprising an amino acid sequence selected from the amino acid sequences of (b), in which one or more amino acids are substituted, deleted, inserted, and/or added, wherein said protein is functionally equivalent to the protein comprising said amino acid sequence selected from the amino acid sequences of (b);
(d) a polynucleotide that hybridizes with a polynucleotide comprising a nucleotide sequence selected from the nucleotide sequences of (a), and that comprises a nucleotide sequence encoding a protein functionally equivalent to the protein encoded by the nucleotide sequence selected from the nucleotide sequences of (a);
(e) a polynucleotide comprising a nucleotide sequence encoding a partial amino acid sequence of a protein encoded by the polynucleotide of (a) to (d);
(f) a polynucleotide comprising a nucleotide sequence with at least 70% identity to the nucleotide sequence of (a).

9. A substantially pure protein encoded by the polynucleotide of claim 8.

10. An antibody against the protein or peptide of any one of claims 6, 7, and 9.

11. A vector comprising the polynucleotide of claim 5 or 8.

12. A transformant carrying the polynucleotide of claim 5 or 8, or the vector of claim 11.

13. A transformant expressively carrying the polynucleotide of claim 5 or 8, or the vector of claim 11.

14. A method for producing the protein or peptide of any one of claims 6, 7, and 9, comprising culturing the transformant of claim 13 and recovering the expression product.

15. An oligonucleotide comprising the nucleotide sequence of claim 8 (a) or the nucleotide sequence complementary to the complementary strand thereof, wherein said oligonucleotide comprises 15 nucleotides or more.

16. Use of the oligonucleotide of claim 15 as a primer for synthesizing a polynucleotide.

17. Use of the oligonucleotide of claim 15 as a probe for detecting a gene.

18. An antisense polynucleotide against the polynucleotide of claim 8, or the portion thereof.

19. A method for synthesizing a polynucleotide, the method comprising:

a) synthesizing a complementary strand using a cDNA library as a template, and using the primer set of claim 2 or 3, or the primer of claim 16; and
b) recovering the synthesized product.

20. The method of claim 19, wherein the cDNA library is obtainable by oligo-capping method.

21. The method of claim 19, wherein the complementary strand is obtainable by PCR.

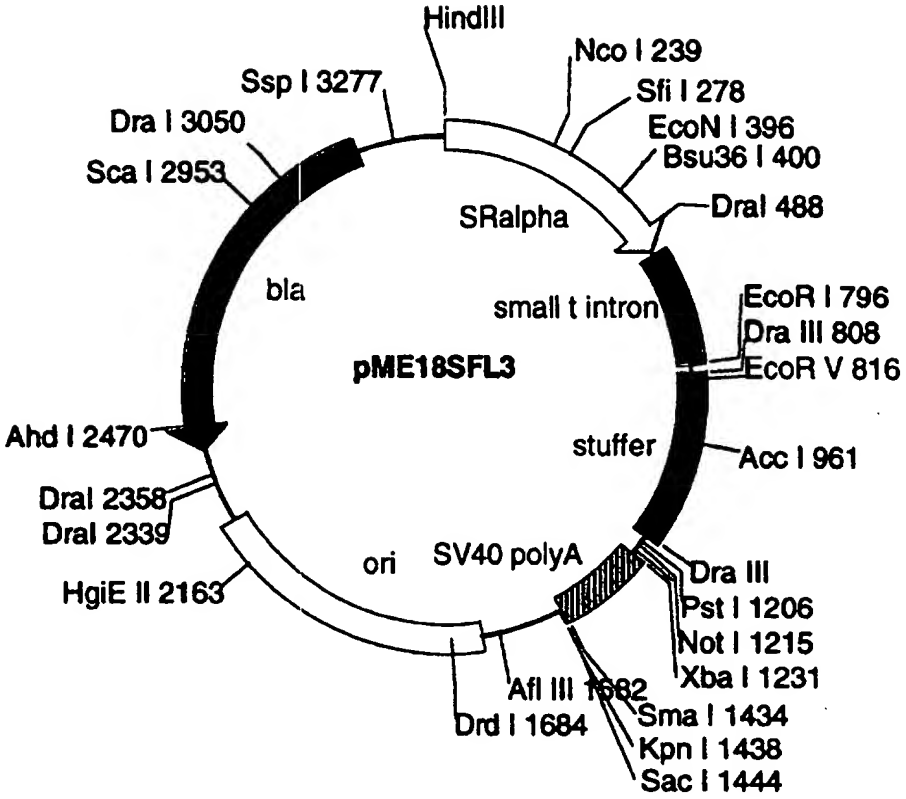
22. A method for detecting the polynucleotide of claim 8, the method comprising:

a) incubating a target polynucleotide with the oligonucleotide of claim 15 under the conditions where hybridization occurs, and
b) detecting the hybridization of the target polynucleotide with the oligonucleotide of claim 15.

23. A database of polynucleotides and/or proteins, the database comprising information on at least one sequence selected from the nucleotide sequences of claim 8 (a) and/or the amino acid sequences of claim 8 (b), or a medium

on which the database is stored.

Figure 1


HindIII
Nco I 239
Sfi I 278
EcoN I 396
Bsu36 I 400
DraI 488
SRalpha
small t intron
EcoR I 796
Dra III 808
EcoR V 816
Acc I 961
stuffer
SV40 polyA
Dra III
Pst I 1206
Not I 1215
Xba I 1231
Sma I 1434
Kpn I 1438
Sac I 1444
Afl III 1682
Drd I 1684
ori
HgiE II 2163
DraI 2339
DraI 2358
Ahd I 2470
bla
Sca I 2953
Dra I 3050
Ssp I 3277
pME18SFL3

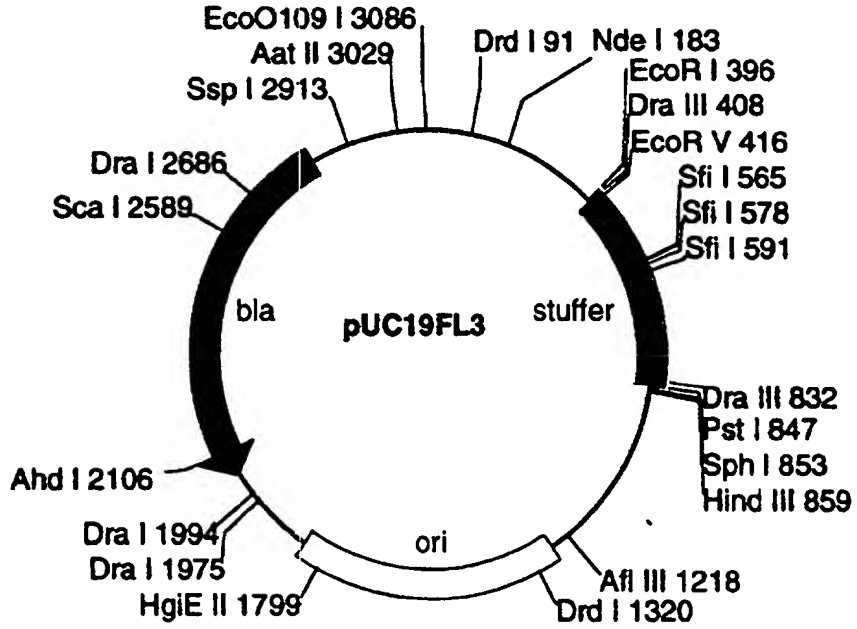

EcoO109 I 3086
Aat II 3029
Ssp I 2913
Dra I 2686
Sca I 2589
Drd I 91
Nde I 183
EcoR I 396
Dra III 408
EcoR V 416
Sfi I 565
Sfi I 578
Sfi I 591
stuffer
Dra III 832
Pst I 847
Sph I 853
Hind III 859
Afl III 1218
Drd I 1320
ori
HgiE II 1799
Dra I 1975
Dra I 1994
Ahd I 2106
bla
pUC19FL3

Figure 2

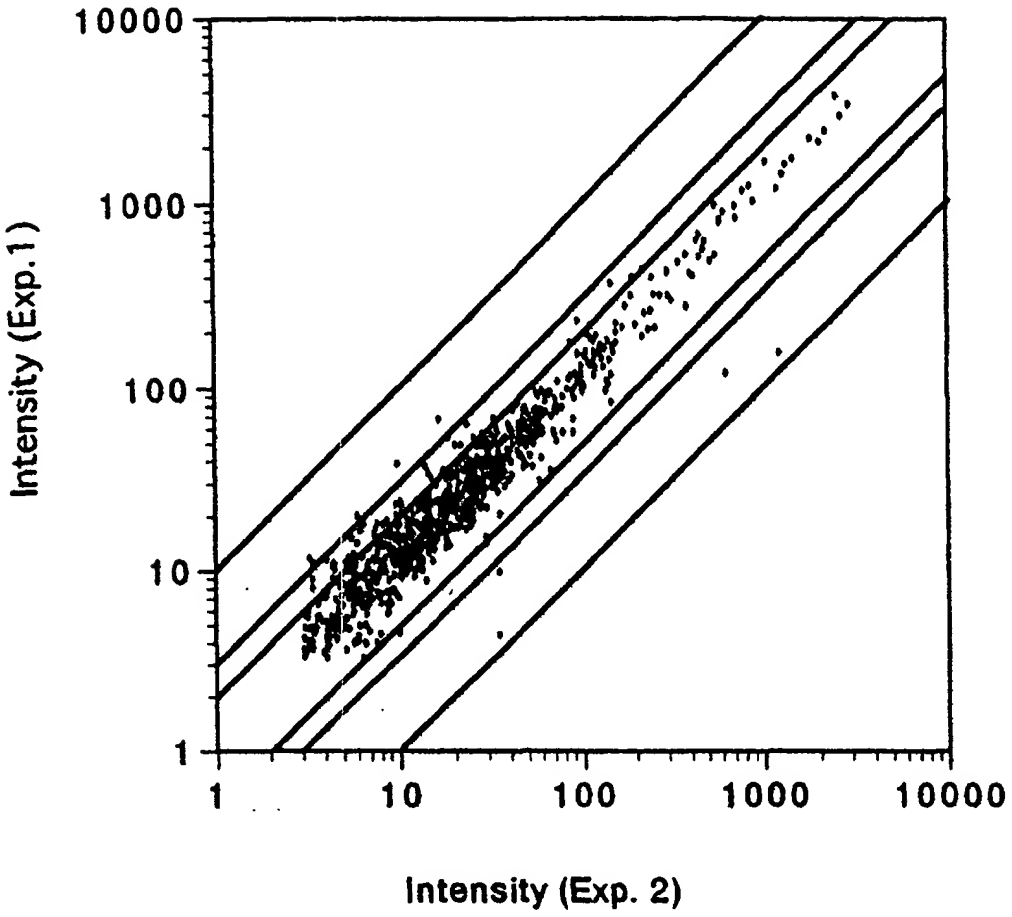
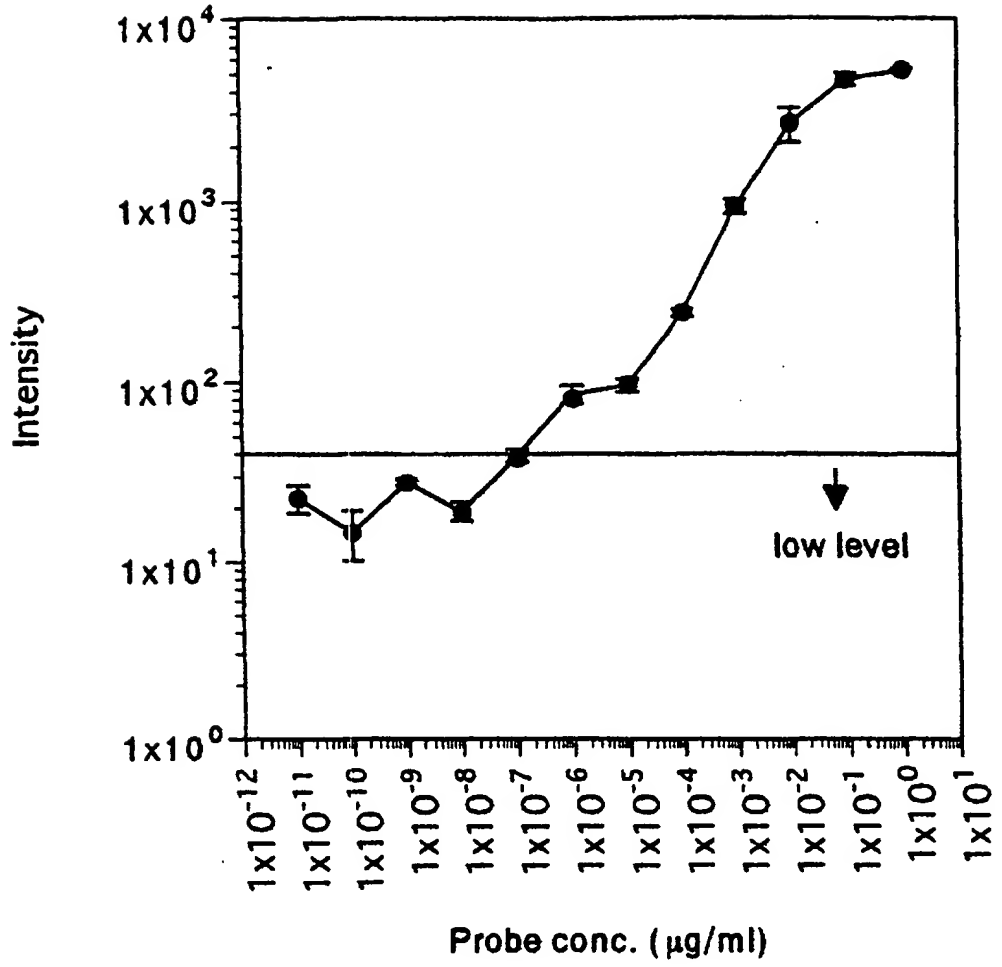

10000
1000
100
10
1
Intensity (Exp.1)
1
10
100
1000
10000
Intensity (Exp. 2)

Figure 3


1x10^4
1x10^3
1x10^2
1x10^1
1x10^0
Intensity
low level
1x10^-12
1x10^-11
1x10^-10
1x10^-9
1x10^-8
1x10^-7
1x10^-6
1x10^-5
1x10^-4
1x10^-3
1x10^-2
1x10^-1
1x10^0
1x10^1
Probe conc. (µg/ml)